

Doped Semiconductors: Role of Disorder

Yuri M. Galperin

*Lectures at Lund University
October-November 1999*

Phone: +46 222 90 77, E-mail: yurigteorfys.lu.se

Permanent address:

*Department of Physics, P.O. Box 1048 Blindern, 0316 Oslo
Phone: +47 22 85 64 95, E-mail: iouri.galperinefys.uio.no*

Contents

I	Lightly Doped Semiconductors	1
1	Isolated impurity states	3
1.1	Shallow Impurities	3
2	Localization of electronic states	11
2.1	Narrow bands and Mott transition	11
2.2	Anderson transition	12
2.2.1	Two-state model	13
2.3	Modern theory of Anderson localization	16
2.3.1	Weak localization	16
2.3.2	Weak localization .. magnetic field	19
2.3.3	Aharonov-Bohm effect	21
2.3.4	$e - e$ interaction	22
2.3.5	Scaling theory	24
3	Impurity band for lightly doped semiconductors.	29
3.1	Low degree of compensation	30
3.2	High degrees of compensation	34
3.3	Specific features of the two-dimensional case	35
4	DC hopping conductance ...	39
4.1	Basic experimental facts	39
4.2	The Abrahams-Miller resistor network model	41
4.2.1	Derivation	41
4.2.2	Conventional treatment	45
5	Percolation theory	47
5.1	Lattice problems	47
5.1.1	“Liquid flow” definition	47
5.1.2	Cluster statistics	48
5.1.3	Percolation through a finite lattice	51
5.1.4	Summary	52
5.2	Continuous problems	53

5.3	Random site problems	56
5.4	Random networks.Infinite cluster topology	58
5.5	Conductivity of strongly inhomogeneous media	60
6	Hopping conductance	65
6.1	Dependence on impurity concentrations	65
6.2	The contribution ρ_3 as a function of impurity concentrations	65
6.3	Activation energy	66
6.3.1	Low degree of compensation	66
6.3.2	High degree of compensation	67
6.4	Variable range hopping	68
6.5	Coulomb gap in the density of states	69
7	AC conductance due to localized states	73
II	Heavily Doped Semiconductors	81
8	Interband light absorption	83

Part I

Lightly Doped Semiconductors

Chapter 1

Isolated impurity states

1.1 Shallow Impurities

A typical energy level diagram is shown on Fig. 1.1 Shallow levels allow a universal description because the spread of wave function is large and the potential can be treated as from the point charge,

$$U(r) = e^2/\kappa r.$$

To find impurity states one has to treat Schrödinger equation (SE) including periodic potential + Coulomb potential of the defect.

Extremum at the center of BZ

Then for small k we have

$$E_n(k) = \frac{\hbar^2 k^2}{2m}.$$

We look for solution of the SE

$$(\mathcal{H}_0 + U)\psi = E\psi$$

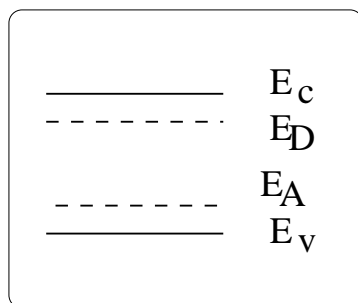


Figure 1.1: band diagram of a semiconductor

in the form

$$\psi = \sum_{n'\mathbf{k}'} B_{n'}(\mathbf{k}') \phi_{n'\mathbf{k}'}(\mathbf{r}),$$

where $\phi_{n'\mathbf{k}'}(\mathbf{r})$ are Bloch states. By a usual procedure (multiplication by $\phi_{n\mathbf{k}}^*(\mathbf{r})$ and integration over \mathbf{r}) we get the equation

$$\begin{aligned} [E_n(\mathbf{k}) - E] B_n(\mathbf{k}) + \sum_{n'\mathbf{k}'} U_{n'\mathbf{k}'}^{n\mathbf{k}} B_{n'}(\mathbf{k}) &= 0 \\ U_{n'\mathbf{k}'}^{n\mathbf{k}} &= \frac{1}{\mathcal{V}} \int u_{n\mathbf{k}}^* u_{n'\mathbf{k}'} e^{i(\mathbf{k}' - \mathbf{k})\mathbf{r}} U(\mathbf{r}) d\mathbf{r}. \end{aligned}$$

Then, it is natural to assume that $B(\mathbf{k})$ is nonvanishing only near the BZ center, and to replace central cell functions u by their values at $\mathbf{k} = 0$. These function rapidly oscillate within the cell while the rest varies slowly. Then within each cell

$$\int_{\text{cell}} u_{n0}^* u_{n'0} d\mathbf{r} = \delta_{nn'}$$

because Bloch functions are orthonormal. Thus,

$$\begin{aligned} [E_n(\mathbf{k}) - E] B_n(\mathbf{k}) + \sum_{n'} U(\mathbf{k}\mathbf{k}') B_n(\mathbf{k}') &= 0 \\ U(\mathbf{k}\mathbf{k}') &= \frac{1}{\mathcal{V}} \int e^{i(\mathbf{k} - \mathbf{k}')\mathbf{r}} U(\mathbf{r}) d\mathbf{r} = -\frac{4\pi e^2}{\kappa \mathcal{V} |\mathbf{k} - \mathbf{k}'|^2}. \end{aligned}$$

Finally we get

$$\left[\frac{\hbar^2 k^2}{2m} - E \right] B_n(\mathbf{k}) - \frac{4\pi e^2}{\kappa \mathcal{V}} \sum_{\mathbf{k}'} \frac{1}{|\mathbf{k} - \mathbf{k}'|^2} B_n(\mathbf{k}')$$

where one can integrate over \mathbf{k} in the infinite region (because $B_n(\mathbf{k})$ decays rapidly).

Coming back to the real space and introducing

$$F(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} \sum_{\mathbf{k}} B_n(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}}$$

we come to the SE for a hydrogen atom,

$$\left[-\frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{\kappa r} \right] F(r) = E F(r).$$

Here

$$\begin{aligned} E_t &= -\frac{1}{t^2} \frac{e^4 m}{2\kappa^2 \hbar^2}, \quad t = 1, 2, \dots \\ F(r) &= (\pi a^3)^{-1/2} \exp(-r/a), \quad a = \hbar^2 \kappa / m e^2. \end{aligned}$$

For the total wave function one can easily obtain

$$\psi = u_{n0}(\mathbf{r}) F(\mathbf{r}).$$

The results are summarized in the table.

Material	κ	m/m_0	E_{1s} (th.) (meV)	E_{1s} (exp.) (meV)
GaAs	12.5	0.066	5.67	Ge: 6.1 Si: 5.8 Se: 5.9 S: 6.1 S: 5.9
InP	12.6	0.08	6.8	7.28
CdTe	10	0.1	13	13.*

Table 1.1: Characteristics of the impurity centers.

Several equivalent extrema

Let us consider silicon for example. The conduction band minimum is located at $k_{z0} = 0.85(2\pi/a)$ in the [100] direction, the constant energy surfaces are ellipsoids of revolution around [100]. There must be 6 equivalent ellipsoids according to cubic symmetry. For a given ellipsoid,

$$E = \frac{\hbar^2}{2m_\ell}(k_z - k_z^0)^2 + \frac{\hbar^2}{2m_t}(k_x^2 + k_y^2).$$

Here $m_\ell = 0.916 m_0$, $m_t = 0.19 m_0$. According to the effective mass theory, the energy levels are N -fold degenerate, where n is the number of equivalent ellipsoids. In real situation, these levels are split due to short-range corrections to the potential. These corrections provide inter-extrema matrix elements. The results for an arbitrary ratio $\gamma = m_t/m_\ell$ can be obtained only numerically by a variational method (Kohn and Luttinger). The trial function was chosen in the form

$$F = (\pi a_{\parallel} a_{\perp}^2)^{-1/2} \exp \left\{ - \left[\frac{x^2 + y^2}{a_{\perp}^2} + \frac{z^2}{a_{\parallel}} \right]^{1/2} \right\},$$

and the parameters a_i were chosen to minimize the energy at given γ . Excited states are calculated in a similar way. The energies are listed in table 1.1.

Material	E_{1s} (meV)			E_{2p_0} (meV)
Si (theor.)		31.27		11.51
Si(P)	45.5	33.9	32.6	11.45
Si(As)	53.7	32.6	31.2	11.49
Si(Sb)	42.7	32.9	30.6	11.52
Ge(theor/)		9.81		4.74
Ge(P)	12.9	9.9		4.75
Ge(As)	14.17	10.0		4.75
Ge(Sb)	10.32	10.0		4.7

Table 1.2: Donor ionization energies in Ge and Si. Experimental values are different because of chemical shift

Impurity levels near the point of degeneracy

Degeneracy means that there are $t > 1$ functions,

$$\phi_{n\mathbf{k}}^j, \quad j = 1, 2, \dots, t$$

which satisfy Schrödinger equation without an impurity. In this case (remember, $\mathbf{k} \approx 0$),

$$\psi = \sum_{j=1}^t F_j(\mathbf{r}) \phi_{n0}^j(br).$$

The functions F_j satisfy matrix equation,

$$\sum_{j'=1}^t \left[\sum_{\alpha, \beta=1}^3 H_{jj'}^{\alpha\beta} \hat{p}_\alpha \hat{p}_\beta + U(\mathbf{r}) \delta_{jj'} \right] F_{j'} = E F_j. \quad (1.1)$$

If we want to include spin-orbital interaction we have to add

$$\mathcal{H}_{so} = \frac{1}{4m_0 c^2} [\mathbf{s} \times \nabla V] \cdot \hat{\mathbf{p}}.$$

Here \mathbf{s} is the spin operator while V is periodic potential. In general \mathcal{H} -matrix is complicated. Here we use the opportunity to introduce a simplified (the so-called invariant) method based just upon the symmetry.

For simplicity, let us start with the situation when spin-orbit interaction is very large, and split-off mode is very far. Then we have 4-fold degenerate system. Mathematically, it can be represented by a pseudo-spin 3/2 characterized by a pseudo-vector \mathbf{J} .

There are only 2 invariants quadratic in \mathbf{p} , namely $\hat{\mathbf{p}}^2$ and $(\hat{\mathbf{p}} \cdot \mathbf{J})^2$. Thus we have only two independent parameters, and traditionally the Hamiltonian is written as

$$\mathcal{H} = \frac{1}{m_0} \left[\frac{\hat{\mathbf{p}}^2}{2} \left(\gamma_1 + \frac{5}{2} \gamma_2 \right) - \gamma_3 (\hat{\mathbf{p}} \cdot \mathbf{J})^2 \right]. \quad (1.2)$$

That would be OK for spherical symmetry, while for cubic symmetry one has one more invariant, $\sum_i \hat{p}_i^2 J_i^2$. As a result, the Hamiltonian is traditionally expressed as

$$\mathcal{H} = \frac{1}{m_0} \left[\frac{\hat{\mathbf{p}}^2}{2} \left(\gamma_1 + \frac{5}{2} \gamma_2 \right) - \gamma_3 (\hat{\mathbf{p}} \cdot \mathbf{J})^2 + (\gamma_3 - \gamma_2) \sum_i \hat{p}_i^2 J_i^2 \right]. \quad (1.3)$$

This is the famous Luttinger Hamiltonian. Note that if the lattice has no inversion center there also linear in \mathbf{p} terms.

Now we left with 4 coupled Schrödinger equations (1.1). To check the situation, let us first put $U(\mathbf{r}) = 0$ and look for solution in the form

$$F_j = A_j(\mathbf{k}/k) e^{i\mathbf{k}\mathbf{r}}, \quad k \equiv |\mathbf{k}|.$$

The corresponding matrix elements can be obtained by substitution $\hbar\mathbf{k}$ instead of the operator \hat{p} into Luttinger Hamiltonian. The Hamiltonian (1.2) does not depend on the direction of \mathbf{k} . Thus let us direct \mathbf{k} along z axis and use representation with diagonal J_z^2 . Thus the system is decoupled into 4 independent equation with two different eigenvalues,

$$E_\ell = \frac{\gamma_1 + 2\gamma}{2m_0} \hbar^2 k^2, \quad E_\ell = \frac{\gamma_1 - 2\gamma}{2m_0} \hbar^2 k^2.$$

If

$$\gamma_1 \pm 2\gamma > 0$$

both energies are positive (here the energy is counted inside the valence band) and called the light and heavy holes. The effective masses are

$$m_{\ell(h)} = m_0 / (\gamma_1 \pm \gamma).$$

The calculations for the full Luttinger Hamiltonian (1.3) require explicit form of \mathbf{J} -matrices. Its solutions lead to the anisotropic dispersion law

$$E_{\ell,h} = \frac{\hbar^2}{2m_0} \left\{ \gamma_1 k^2 \pm 4 \left[\gamma_2^2 k^4 + 12(\gamma_3^2 - \gamma_2^2)(k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2) \right]^{1/2} \right\}.$$

The parameters of Ge and Si are given in the Table 1.1

Material	γ_1	γ_2	γ_3	Δ	κ
Ge	4.22	0.39	1.44	0.044	11.4
Si	13.35	4.25	5.69	0.29	15.4

Table 1.3: Parameters of the Luttinger Hamiltonian for Ge and Si

The usual way to calculate acceptor states is variational calculation under the spherical model (1.2). In this case the set of 4 differential equations can be reduced to a system of 2 differential equation containing only 1 parameter, $\beta = m_\ell/m_h$.

Asymptotic behavior of the impurity-state wave function

At large distances the wave functions of the localized states decrease exponentially, and for many problems it is sufficient to know only the exponential factor. Let us restrict ourselves by the effective mass method which is valid at the distances much larger than the lattice spacing. For the case of non-degenerate band the Schrödinger equation for the envelope function, F , is of the form

$$[\mathcal{H}(-i\hbar\nabla) + U(\mathbf{r})]F = EF. \quad (1.4)$$

The asymptotic behavior can be understood within quasiclassical approximation,

$$F = \exp[iS(\mathbf{r})/\hbar]$$

where deep under the barriers the imaginary part of the action S is large.¹ Using the above form for F and following the rules of quasiclassical approach where only the first derivatives of S should be kept we get a classical Hamilton-Yacoby equation

$$\mathcal{H}(\nabla S) + U(\mathbf{r}) = 0.$$

Now let us assume that we know both the action $S(\mathbf{r})$ and its gradient ∇S at a surface σ . Now we can draw a ray through each point of σ which satisfies classical equations of

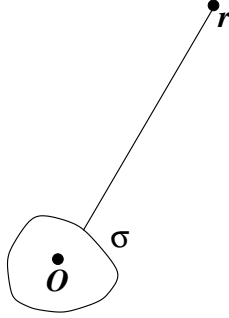


Figure 1.2: A quasi-classical trajectory passing through the point .

motion and has $\mathbf{p} = \nabla S$ on the surface σ . Then at arbitrary point \mathbf{r} we have

$$S(\mathbf{r}) = S(\mathbf{r}_\sigma) + \int_{\mathbf{r}_\sigma}^{\mathbf{r}} \mathbf{p} \cdot d\mathbf{r}'.$$

According to classical mechanics, we have the Hamilton equations

$$\dot{\mathbf{r}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}} = -\frac{\partial U}{\partial \mathbf{r}}. \quad (1.5)$$

Since by definition $S(\mathbf{r}_\sigma) \ll S(\mathbf{r})$, and outside the range of the potential $\mathbf{p} = \mathbf{p}_0 = \text{const}$,

$$S(\mathbf{r}) = \mathbf{p}_0 \cdot \mathbf{r}.$$

The group velocity $\mathbf{v} = \partial E / \partial \mathbf{p}$ corresponding to the vector \mathbf{p}_0 should be perpendicular to the surface σ at the incident point, while its absolute value is determined by the condition

$$\mathcal{H}(\mathbf{p}_0) = E. \quad (1.6)$$

As a result, at large distances the wave function is a plane wave

$$F = e^{i\mathbf{p}_0 \cdot \mathbf{r} / \hbar}.$$

¹The condition for the validity of the quasiclassical approach reads

$$\left| \frac{d}{d\mathbf{r}} \frac{\hbar}{|dS/d\mathbf{r}|} \right| \ll 1.$$

Consider the simplest case, $\mathcal{H}(\mathbf{p}) = p^2/2m$ where the group velocity $\mathbf{v} \equiv \nabla_{\mathbf{p}}\mathcal{H} \parallel \mathbf{p}$. Then for *negative* energies we can write

$$F(r) = e^{-qr}, \quad q = \frac{\sqrt{2m|E|}}{\hbar}.$$

This solution appears *exact* for the hydrogen-like spectrum with

$$q = a^{-1}, \quad a = \hbar^2 \kappa / m e^2.$$

The quantity a is the effective Bohr radius. Remember that for the validity of the approximation under consideration the inequality

$$qa_0 \ll 1, \quad a_0 = \hbar^2 / m_0 e^2$$

must be met.

Now let us consider the case of ellipsoidal spectrum,

$$\mathcal{H}(\mathbf{p}) = \frac{p_x^2 + p_y^2}{2m_l} + \frac{p_z^2}{2m_t}. \quad (1.7)$$

Introducing the mass ratio $\gamma \equiv m_t/m_l$ and the quantity $d = \sqrt{p_{0x}^2 + p_{0y}^2 + \gamma^2 p_{0z}^2}$ we have the set of equations

$$p_{0x} = n_x d, \quad p_{0y} = n_y d, \quad p_{0z} = n_z d / \gamma, \quad (1.8)$$

Substituting these relations to Eq. (1.6) for negative E we get

$$d^2 = \frac{2m_t |E|}{n_x^2 + n_y^2 + n_z^2 / \gamma}.$$

As a result,

$$q(\mathbf{n}) = \hbar^{-1} \sqrt{2m_t |E| (n_x^2 + n_y^2 + n_z^2 / \gamma)}. \quad (1.9)$$

We observe that the shape of the wave function is also ellipsoidal. In particular, for Ge we have $q_l^{-1} = 13.8 \text{ \AA}$ ($n_x = n_y = 0$), $q_t^{-1} = 61.4 \text{ \AA}$ ($n_z = 0$).

As we have mentioned, in semiconductors with several ellipsoids (Ge, Si) the wave function is represented by a linear combination of the functions corresponding to different ellipsoids with different $q_j(\mathbf{n})$. At very large distances we are interested in

$$q(\mathbf{n}) = \min\{q_j(\mathbf{n})\}.$$

We do not discuss here the case of degenerate bands. One can find details in the book by Shklovskii and Efros [1]

Chapter 2

Localization of electronic states

What happens when the number of doping impurities is large and their wave functions overlap?

2.1 Narrow bands and Mott transition

As a simple example, consider an ordered lattice of impurities, the potential being

$$V(\mathbf{r}) = \sum_j U(\mathbf{r} - \mathbf{r}_j) .$$

Assume that we know the eigenstates ϕ_n and eigenvalues, E_n , of a single-impurity problem. Here we shall use single-band approximation, and therefore ignore Bloch (central-cell) factors. Also, we assume the impurity bandwidth is less than the spacing between E_n and restrict ourselves by the lowest level, E_0 . Along the tight-binding method, we can expand the wave functions in terms of the above eigenvalues,

$$\psi = \sum_j a_j \phi(\mathbf{r} - \mathbf{r}_j), \quad \sum_j |a_j|^2 = 1 ,$$

the energy being

$$E = \sum_{\mathbf{m}} e^{i\mathbf{k}\mathbf{m}} I(\mathbf{m}) .$$

The only important property of the *overlap integrals*, $I(m)$ is that it is small (tight-binding approximation!).

In this way we get energy bands. For example, for a simple cubic lattice

$$E = 2I(b_0) \sum_i \cos(k_i b_0)$$

where here b_0 denotes the lattice constant for impurity lattice. Expanding near $\mathbf{k} = 0$ we get

$$E(k) = 6I(b_0) - I(b_0)k^2 b_0^2$$

the effective mass being $m^* = \hbar^2/2Ib_0^2$. As the lattice constant increases, $I(b_0) \propto \exp(-\beta b_0/a)$ where β is a number of the order 1.

According to our model, each impurity adds one electron, and each state possesses 2-fold spin degeneracy. Thus the impurity lattice is a *metal*.

Is this correct? In principle, no, because we neglected electron-electron interaction. In fact, two electrons at the same site repel each other the energy can be estimated as $U_0 \approx e^2/a$. If U_0 is comparable with the bandwidth $\sim I(b_0)$, then one must allow for the interaction. At large b_0 the band is narrow and this is the case. Let us plot the electron bands as a function of $1/b_0$. The insulator-to-metal transition of this kind is usually called

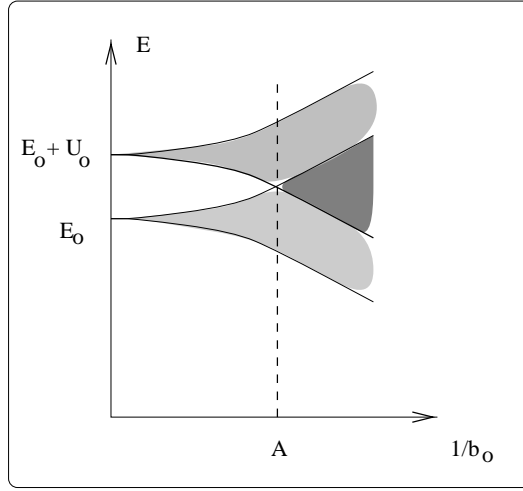


Figure 2.1: Dependence of electron bands on impurity sublattice period b_0 . To the left of point A is an insulator, to the right a metal.

the Mott transition.

2.2 Anderson transition

We come back to single-electron approximation and concentrate on disorder. Suppose that impurities are still ordered, but depths of the potential wells are random, Fig. (2.2) The effective Hamiltonian has the form

$$\mathcal{H} = \sum_j \epsilon_j a_j^\dagger a_j + \sum_{m,j \neq 0} I(m) a_j^\dagger a_{j+m}.$$

The energies ϵ are assumed to be uncorrelated and random, the distribution being

$$P(\epsilon) = \begin{cases} 1/A & , \quad |\epsilon| < A/2 \\ 0 & , \quad |\epsilon| > A/2 \end{cases}$$

Anderson has formulated the following criterion for localization. Let us place a particle at the site i and study its decay in time due to quantum smearing of the wave packet. The state is called extended if $|\psi_i(t)|^2 \rightarrow 0$ at $t \rightarrow \infty$.

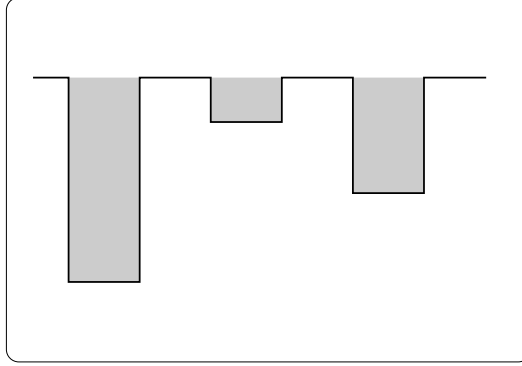


Figure 2.2: Potential wells in the Anderson model.

The *Anderson model* has one dimensionless parameter, A/I , where I is the overlap integral for nearest neighbors. In a 3D system there is a critical value, A_c/I , at which delocalized states begin to appear in the middle band. In 1D case the states are localized *at any disorder*. 2D case is a marginal, the states being also localized at any disorder.

2.2.1 Two-state model

Now let us turn to 3D case and try to discuss Anderson result in a simplified form. Namely, consider two potential wells. When isolated, they can be characterized by the energies ϵ_1 and ϵ_2 and wave functions ϕ_1 and ϕ_2 . At $\epsilon_1 = \epsilon_2$ the wave functions are

$$\Psi_{I,II} = \frac{1}{\sqrt{2}}(\phi_1 \pm \phi_2),$$

the energy difference being $E_I - E_{II} = 2I$. This solution is reasonably good at $|\epsilon_1 - \epsilon_2| \ll I$. In general,

$$\psi_{I,II} = c_1\phi_1 \pm c_2\phi_2,$$

and we have a matrix Schrödinger equation,

$$\begin{pmatrix} \Delta/2 - E & I \\ I^* & -\Delta/2 - E \end{pmatrix}.$$

Here the origin for energy is taken at $(\epsilon_1 + \epsilon_2)/2$, while $\Delta \equiv \epsilon_1 - \epsilon_2$. The secular equation is thus

$$E^2 - (\Delta/2)^2 - |I|^2 = 0 \rightarrow E_{I,II} = \pm \sqrt{(\Delta/2)^2 + |I|^2}.$$

Consequently,

$$E_I - E_{II} = \sqrt{\Delta^2 + 4|I|^2}. \quad (2.1)$$

The ratio

$$\frac{c_1}{c_2} = \frac{I}{\Delta/2 \pm \sqrt{(\Delta/2)^2 + |I|^2}}.$$

Thus at $\Delta \gg I$ either c_1 or c_2 is close to 1, and collectivization does not occur.

Now, following Thouless, let us chose a band

$$\delta/2 \leq \epsilon \leq \delta/2, \quad \delta \sim I$$

and call the sites *resonant* if their energy falls into the band. Then let us look for *connected* resonant states which share a site. Non-resonant sites can be disregarded as $I \ll A$.

It is clear that it must be a threshold in the quantity A/I where the transition takes place. If one assumes that the connected cluster is a 1D chain, then the bandwidth is $4I$. In such a model,

$$\frac{A_c}{I} = \frac{4}{x_c}$$

where x_c is the *percolation threshold* for the site problem. This is quite a good estimate, see the Table 2.1 Finally, one arrives at the following profile of density-of-states (DOS)

Lattice	x_c	$4/x_c$	A_c/I
Diamond	0.43	9.3	8.0
Simple cubic	0.31	12.9	15

Table 2.1: Percolation thresholds and critical values A_c/I obtained numerically for different lattices.

for the Anderson model, Fig. 2.3 According to this picture, there both extended and

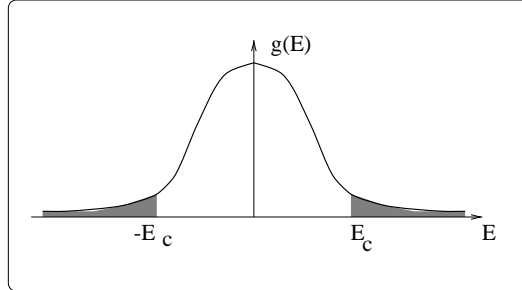


Figure 2.3: Density of states in the Anderson model. Localized states are shaded. Mobility edges are denoted as E_c .

localized states separated by the *mobility edges*. If we vary the electron density the Fermi level will move with respect to the bands and it may cross the mobility edge. At this point a transition in conductance takes place which is conventionally called *the Anderson transition*.

To illustrate the Anderson transition let us discuss the case of an amorphous semiconductor. The presence of a short-range order preserves the band picture, so there are forbidden gaps which separate the allowed regions of energy. However, structural “faults” give rise to the *tail* in the DOS, and the band boundaries appear smeared. If the Fermi

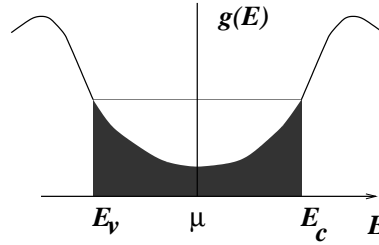


Figure 2.4: DOS in an amorphous semiconductor. Localized states are shadowed.

level falls into the localized region the transport is possible only due to thermal activation into the delocalized states region, or by hopping between the localized states. Both mechanisms lead to an *exponential* temperature dependence of conductance at low temperatures. At very low temperatures a weak tunneling between the localized states is the only source of transport, so it becomes essentially temperature-independent.

If we increase the electron concentration, the Fermi level moves to the direction of delocalized states, and the insulator-to metal-transition takes place. As a result, a temperature-activated hopping crosses over to a weak temperature dependence which is called the *metallic conductance*. Such transition is called the *Mott-Anderson* transition because it bears the features of both models.

Anderson localization appears important for metal-insulator-semiconductor (MIS) structures which are extensively used in modern electronics. Such a structure is most often realized with a SiO_2 film (insulator) sandwiched between Si substrate and a flat metal electrode. Electrostatic forces force the energy bands to bend to keep a constant electrochemical potential. As a result, the charges re-distribute creating on the semiconductor surface a narrow *inversion layer*, i. e. the layer with the carriers of the type opposite to the ones in the bulk, see Fig. 2.5. By changing the applied potential one can control con-

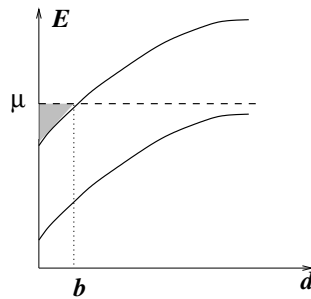


Figure 2.5: Band diagram near the MIS interface. The inversion layer is shadowed. μ is the chemical potential.

ductivity of the inversion layer over a wide range. This principle is used for the so-called field-effect transistors (FET). At low carrier concentrations the conductance of the layer is activated. As the concentration increases, the Fermi level μ moves towards the extended states, the conductivity increases and becomes a weak function of the temperature.

Now a question arises: Is the Mott-Anderson transition *abrupt* or *continuous*? The first answer to this question has been suggested by Mott who arrived at the concept of *minimal metallic conductivity*. This concept is close to the so-called *Ioffe-Regel principle* that the mean free path ℓ which enters the Drude formula for 3 dimensional electron gas (3DEG),

$$\sigma = \frac{e^2 p_F^2 \ell}{3\pi^2 \hbar}$$

cannot be less than the de Broglie electron wavelength, \hbar/p_F . In this way we obtain

$$\sigma_{\min} \approx \frac{e^2}{\hbar \ell_{\min}}.$$

Within the Anderson model, $\ell_{\min} = a_0$ where a_0 is the lattice constant, while for the Mott model ℓ_{\min} corresponds to the typical distance between the impurities. According to Mott's suggestion, at $T \rightarrow 0$ the conductivity reaches its minimal value, and then drops to zero, see Fig. 2.6. Mott's arguments for the two-dimensional case lead to the estimate

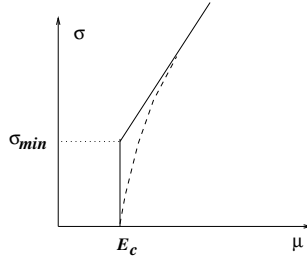


Figure 2.6: Mott's suggestion of the metal-to-insulator transition. Zero-temperature conductivity as a function of the Fermi level μ .

$$\sigma_{\min} \sim \frac{e^2}{\hbar}.$$

The modern theory contradicts to the Mott's concept, it predicts a *continuous* transition.

2.3 Modern theory of Anderson localization

We start from the case of metallic conductance to find in which way disorder influences the conductance. Then we shall discuss the aspects of strong localization.

2.3.1 Weak localization

Consider noninteracting electrons having $p_F \ell \gg \hbar$ and passing between the points A and B through scattering media. The probability is

$$W = \left| \sum_i A_i \right|^2 = \sum_i |A_i|^2 + \sum_{i \neq j} A_i A_j^*. \quad (2.2)$$

Here A_i is the propagation *amplitude* along the path i . The first item – classical probability, the second one – *interference* term.

For the majority of the trajectories the phase gain,

$$\Delta\varphi = \hbar^{-1} \int_A^B \mathbf{p} d\mathbf{l} \gg 1, \quad (2.3)$$

and interference term vanishes. Special case - trajectories with self-crossings. For these

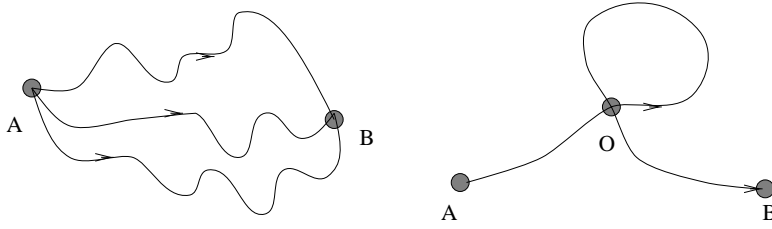


Figure 2.7: Feynman paths responsible for weak localization

parts, if we change the direction of propagation, $\mathbf{p} \rightarrow -\mathbf{p}$, $d\mathbf{l} \rightarrow -d\mathbf{l}$, the phase gains are the same, and

$$|A_1 + A_2|^2 = |A_1|^2 + |A_2|^2 + 2A_1A_2^* = 4|A_1|^2.$$

Thus quantum effects double the result. As a result, the total scattering probability at the scatterer at the site O *increases*. As a result, conductance *decreases* - predecessor of localization.

Probability of self-crossing. The cross-section of the site O is in fact de Broglie electron wave length, $\lambda \sim \hbar/p_F$. Moving diffusively, it covers the distance $\sqrt{x_i^2} \sim \sqrt{Dt}$, covering the volume $(Dt)^{d/2}b^{3-d}$. Here d is the dimensionality of the system, while b is the “thickness” of the sample. To experience interference, the electron must enter the interference volume, the probability being

$$\frac{v\lambda^2 dt}{(Dt)^{d/2}b^{3-d}}.$$

Thus the relative correction is

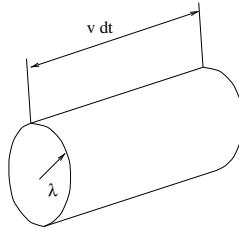


Figure 2.8: On the calculation of the probability of self-crossing.

$$\frac{\Delta\sigma}{\sigma} \sim - \int_{\tau}^{\tau_{\varphi}} \frac{v\lambda^2 dt}{(Dt)^{d/2}b^{3-d}} \quad (2.4)$$

The upper limit is the so-called *phase-breaking time*, τ_φ . Physical meaning – it is the time during which the electron remains coherent. For example, any inelastic or spin-flip scattering gives rise to phase breaking.

3d case

In a 3d case,

$$\frac{\Delta\sigma}{\sigma} \sim -\frac{v\lambda^2}{D^{3/2}} \left(\frac{1}{\sqrt{\tau}} - \frac{1}{\sqrt{\tau_\varphi}} \right) \sim -\left(\frac{\lambda}{\ell} \right)^2 + \frac{\lambda^2}{\ell L_\varphi}. \quad (2.5)$$

Here we have used the notations

$$D \sim v\ell, \quad \tau \sim \ell/v, \quad L_\varphi = \sqrt{D\tau_\varphi}. \quad (2.6)$$

On the other hand, one can rewrite the Drude formula as

$$\sigma \sim \frac{n_e e^2 \tau}{m} \sim \frac{n_e e^2 \ell}{p_F} \sim \frac{e^2 p_F^2 \ell}{\hbar^3}.$$

In this way, one obtains

$$\Delta\sigma \sim -\frac{e^2}{\hbar\ell} + \frac{e^2}{\hbar L_\varphi}. \quad (2.7)$$

Important point: The second item, though small, is of the most interest. Indeed, it is temperature-dependent because of inelastic scattering. There are several important contributions:

- Electron-electron scattering:

$$\tau_\varphi \sim \hbar\epsilon_F/T^2.$$

Thus

$$\frac{\sigma(T) - \sigma(0)}{\sigma(0)} \sim \left(\frac{\lambda}{\ell} \right)^{3/2} \left(\frac{T}{\epsilon_F} \right).$$

It is interesting that at low temperatures this quantum correction can exceed the classical temperature-dependent contribution of e - e -scattering which is $\propto T^2$. It is also important to note that the above estimate is obtained for a clean system. It should be revised for disordered systems where electron-electron interaction appears more important (see later).

- Electron-phonon-interaction. In this case,¹

$$\tau_\varphi \sim \frac{\hbar^2 \omega_D^2}{T^3},$$

and

$$\frac{\sigma(T) - \sigma(0)}{\sigma(0)} \sim \left(\frac{\lambda}{\ell} \right)^{3/2} \left(\frac{T}{\epsilon_F} \right)^{1/2} \left(\frac{T}{\hbar\omega_D} \right).$$

¹Under some limitations.

Thus there is a cross-over in the temperature dependencies. To obtain the dominating contribution one has to compare τ_φ^{-1} . Consequently, at low temperatures e - e -interaction is more important than the e - ph one, the crossover temperature being

$$T_0 \sim (\hbar\omega_D^2/\epsilon_F) \sim 1 \text{ K}.$$

Low-dimensional case

If the thickness b of the sample is small ², then the interference volume $\lambda^2 v dt$ has to be related to $(Dt)^{d/2} b^{3-d}$. For a film $d = 2$, while for a quantum wire $d = 1$. As a result,

$$\frac{\Delta\sigma}{\sigma} \sim -\frac{v\lambda^2}{D} \begin{cases} b^{-1} \ln(\tau_\varphi/\tau), & d = 2 \\ b^{-2} L_\varphi, & d = 1 \end{cases} \quad (2.8)$$

It is convenient to introduce *conductance* as

$$G \equiv \sigma b^{3-d}.$$

We have,

$$\Delta G \sim -\frac{e^2}{\hbar} \begin{cases} \ln(L_\varphi/\ell), & d = 2 \\ L_\varphi, & d = 1 \end{cases} \quad (2.9)$$

Important note: In low-dimensional systems the main mechanism of the phase breaking is different. It is so-called low-momentum-transfer electron-electron interaction which we do not discuss in detail

2.3.2 Weak localization in external magnetic field

In a magnetic field one has to replace $\mathbf{p} \rightarrow \mathbf{p} + (e/c)\mathbf{A}$ (remember - we denote electronic charge as $-e$). Thus the product AA^* acquires an additional phase

$$\Delta\varphi_H = \frac{2e}{c\hbar} \oint \mathbf{A} d\mathbf{l} = \frac{2e}{c\hbar} \oint \text{curl } \mathbf{A} d\mathbf{S} = 4\pi \frac{\Phi}{\Phi_0} \quad (2.10)$$

where Φ is the magnetic flux through the trajectory while $\Phi_0 = 2\pi\hbar c/e$ is the so-called *magnetic flux quantum*. Note that it is 2 times greater than the quantity used in the theory of superconductivity for the flux quantum.

Thus magnetic field behaves as an additional *dephasing mechanism*, it “switches off” the localization correction, increases the conductance. In other words, we observe *negative* magnetoresistance which is very unusual.

To make estimates, introduce the typical dephasing time, t_H , to get $\Delta\varphi_H \sim 2\pi$ for the trajectory with $L \sim \sqrt{Dt_H}$. In this way,

$$t_H \sim \Phi_0/(HD) \sim l_H^2/D, \quad l_H = \sqrt{c\hbar/eH}. \quad (2.11)$$

²The general form of the criterion depends on the relationship between the film thickness, b , and the elastic mean free path ℓ . The result presented is correct at $b \ll \ell$. At $L \gg b \gg \ell$ one can replace the lower limit τ of the integral (2.4) by $\tau_b = b^2/D$.

The role of magnetic field is important at

$$t_H \leq \tau_\varphi \quad \rightarrow \quad H \geq H_0 \sim \Phi_0 / (D\tau_\varphi) \approx \hbar c / L_\varphi^2.$$

Note that at $H \approx H_0$

$$\omega_c \tau \sim \hbar / (\epsilon_F \tau_\varphi) \ll 1$$

that means the absence of classical magnetoresistance. *Quantum effects manifest themselves in extremely weak magnetic fields.*

To get quantitative estimates one has to think more carefully about the geometry of diffusive walks. Let consider the channel of 2DEG with width W . The estimates given above are valid for $\ell, L_\varphi \ll W$, exact formulas could be found, e. g. in Ref. [2]. Usually H_0 is very weak, at $L_\varphi = 1 \mu\text{m}$ $H_0 \approx 1 \text{ mT}$. The suppression of weak localization is complete at $H \geq \hbar / e\ell^2$, still under conditions of classically weak magnetic field, $\omega_c \tau \ll 1$.

The situation is a bit different at $W \leq L_\varphi$, this case can be mentioned as one-dimensional for the problem in question, see Fig. 2.9. Then a typical enclosed area is

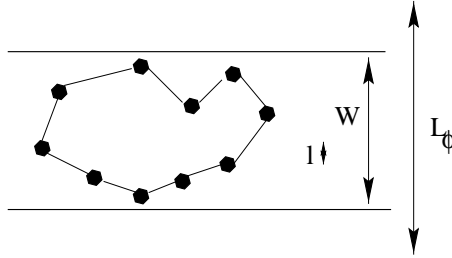


Figure 2.9: On quasi 1D weak localization.

$S \sim WL_\varphi$, and the unit phase shift takes place at

$$t_H \sim L_H^4 / DW^2, \quad \rightarrow \quad H_0 \sim \hbar c / eWL_\varphi.$$

Some experimental results on magnetoresistance of wide and narrow channels are shown in Fig. 2.10.

There are also several specific effects in the interference corrections:

- anisotropy of the effect with respect of the direction of magnetic field (in low-dimensional cases);
- spin-flip scattering acts as a dephasing time;
- oscillations of the longitudinal conductance of a hollow cylinder as a function of magnetic flux. The reason – typical size of the trajectories.

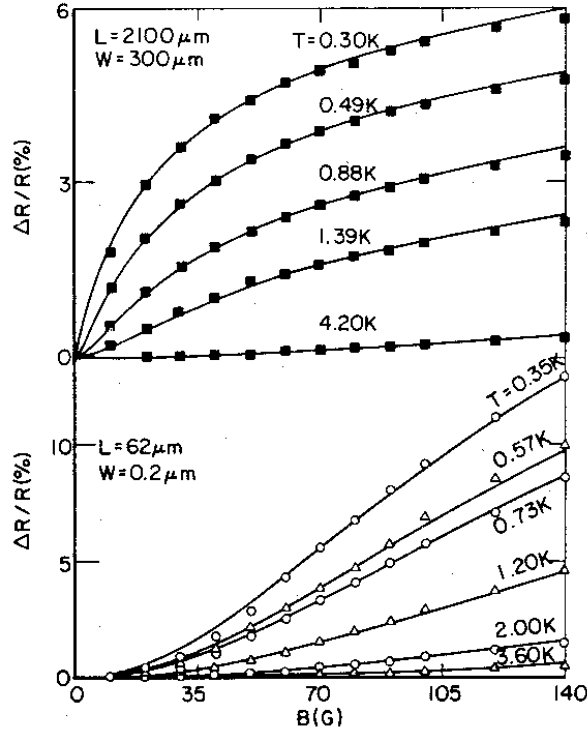


Figure 2.10: Experimental results on magnetoresistance due to 2D weak localization (upper panel) and due to 1D weak localization in a narrow channel (lower panel) at different temperatures. Solid curves are fits based on theoretical results. From K. K. Choi *et al.*, Phys. Rev. B. **36**, 7751 (1987).

2.3.3 Aharonov-Bohm effect

Magnetoconductance corrections are usually aperiodic in magnetic field because the interfering paths includes a continuous range of magnetic flux values. A *ring* geometry, in contrast, encloses a continuous well-defined flux Φ and thus imposes a fundamental periodicity,

$$G(\Phi) = G(\Phi + n\Phi_0), \quad \Phi_0 = 2\pi\hbar c/e.$$

Such a periodicity is a consequence of gauge invariance, as in the originally thought experiment by Aharonov and Bohm (1959). The fundamental periodicity

$$\Delta B = \frac{2\pi\hbar c}{eS}$$

comes from interference of the trajectories which make one half revolution along the ring, see Fig. 2.11. There is an important difference between $\hbar c/e$ and $\hbar c/2e$ oscillations. The first ones are sample-dependent and have random phases. So if the sample has many rings in series or in parallel, then the effect is mostly averaged out. Contrary, the second oscillations originate from time-reversed trajectories. The proper contribution leads to a

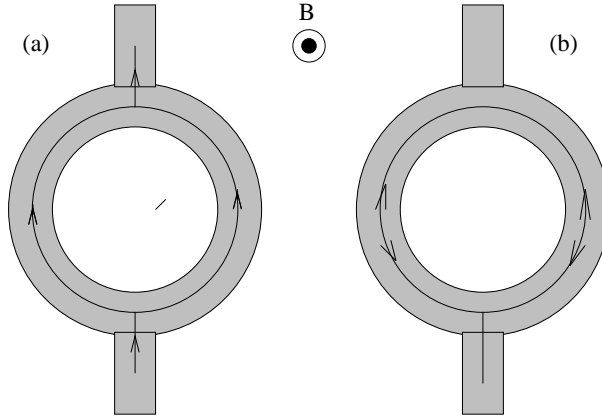


Figure 2.11: Illustration of Aharonov-Bohm effect in a ring geometry. (a) Trajectories responsible for hc/e periodicity, (b) trajectories of the pair of time-reversed states leading to $hc/2e$ -periodicity.

minimum conductance at $H = 0$, thus the oscillations have the same phase. This is why $hc/2e$ -oscillations survive in long hollow cylinders. Their origin is a periodic modulation of the weak localization effect due to coherent backscattering. Aharonov-Bohm oscillations in long hollow cylinders, Fig. 2.12, were predicted by Altshuler, Aronov and Spivak [3] and observed experimentally by Sharvin and Sharvin [4]. A rather simple estimate for the magnitude of those oscillations can be found in the paper by Khmelnitskii [5]. Up to now

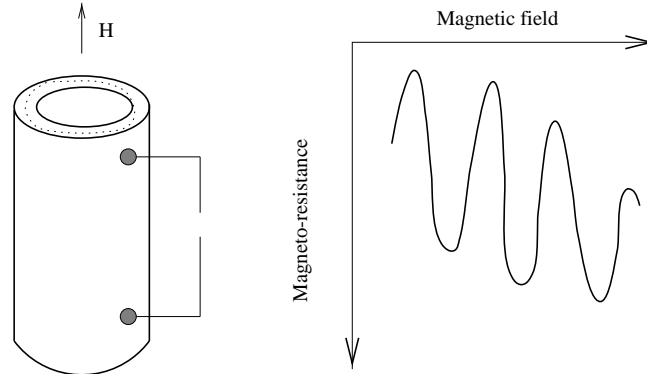
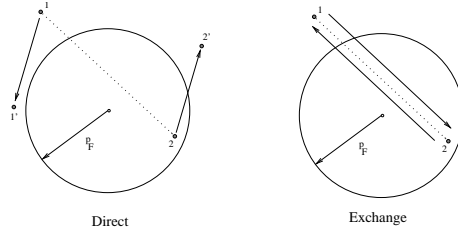


Figure 2.12: On the Aharonov-Bohm oscillations in a long hollow cylinder.

we have analyzed the localization corrections ignoring electron-electron interaction. In the following section we shall discuss the influence of disorder on $e - e$ interaction.

2.3.4 Electron-electron interaction in a weakly disordered regime

Let us discuss the effect of the $e - e$ interaction on the density of states. Let us concentrate

Figure 2.13: On the calculation of $e - e$ interaction.

on the exchange interaction, shown in a right panel of Fig. 2.13

$$\Delta\epsilon = - \int_{|\mathbf{p}-\hbar\mathbf{k}|\leq p_F} g(\mathbf{k}) \frac{d^3k}{(2\pi\hbar)^3}. \quad (2.12)$$

Here $g(\mathbf{k})$ is the Fourier component of the interaction potential, the sign “-” is due to exchange character of the interaction. In the absence of screening $g(k) = 4\pi e^2/k^2$, and

$$\begin{aligned} \Delta\epsilon &= -4\pi e^2 \int_{|\mathbf{p}-\hbar\mathbf{k}|\leq p_F} k^{-2} \frac{d^3k}{(2\pi)^3} \\ &= -4\pi e^2 \hbar^2 \int_{p' < p_F} \frac{1}{|\mathbf{p} - \mathbf{p}'|^2} \frac{d^3p'}{(2\pi\hbar)^3} \\ &= -\frac{e^2}{\pi\hbar} \left(p_F - \frac{p^2 - p_F^2}{2p} \ln \frac{p + p_F}{p - p_F} \right). \end{aligned} \quad (2.13)$$

To obtain this formula it is convenient to use spherical coordinates - $(dp') = 2\pi p'^2 dp' d(\cos\phi)$ and auxiliary integral

$$\int_{-1}^1 \frac{d(\cos\phi)}{p^2 + p'^2 - 2pp'\cos\phi} = \frac{1}{pp'} \ln \left[\frac{p + p'}{p - p'} \right]. \quad (2.14)$$

At small $p - p_F$ it is convenient to introduce $\xi = v(p - p_F) \ll \epsilon_F$ to get (omitting the irrelevant constant)

$$\Delta\epsilon = -(e^2/\pi\hbar v)\xi \ln(2p_F v/\xi). \quad (2.15)$$

Screening can be allowed for by the replacement $k^{-2} \rightarrow (k^2 + \kappa^2)^{-1}$ that replaces $p \pm p'$ in the argument of the logarithm in Eq. (2.14) by $\sqrt{(p \pm p')^2 + (\hbar\kappa)^2}$. As a result,

$$\Delta\epsilon = -(e^2/\pi\hbar v)\xi \ln(2p_F/\hbar\kappa) \quad \text{at} \quad \hbar\kappa \ll p_F. \quad (2.16)$$

This is a simple-minded estimate because it ignores the interference of the states. Indeed, if the two states differ in the energy by $|\xi|$ the coherence time is $\hbar/|\xi|$. If the electron returns back for this time, then the effective interaction constant increases by

$$\alpha_\xi = \int_\tau^{\hbar/|\xi|} \frac{v\lambda^2 dt}{(Dt)^{d/2} b^{3-d}}. \quad (2.17)$$

Thus

$$e_{eff}^e = e^2(1 + \alpha_\xi).$$

In a similar way

$$\frac{\Delta v}{v} = \frac{\Delta g}{g} = \frac{e^2}{\hbar v} \alpha_\xi.$$

Here g is the density of electronic states. Consequently, since $\sigma \propto \nu$, we arrive to a specific correction to conductance. Comparing this correction with the interference one we conclude that interaction dominates in 3d case, has the same order in 2d case and not important in 1d case. Another important feature is that the interaction effects are limited by the coherence time $\hbar/|\xi| \approx \tau_T = \hbar/k_B T$ rather than by τ_φ . Usually, $\tau_\varphi \gg \tau_T$. Consequently, the interference effects can be destroyed by weaker magnetic fields than the interaction ones (important for separation of the effects).

2.3.5 Scaling theory of Anderson localization

If we come back to the interference correction for $d = 2, 1$ we observe that it *increases* with τ_φ , or at $T \rightarrow 0$. Thus the corrections becomes not small. We can also prepare the samples with different values of the mean free path ℓ .

What happens if the corrections are not small? Anderson (1958) suggested localization of electronic states at $T = 0$. This suggestion has been later proved for an infinite 1d system, as well as for an infinite wire of finite thickness (Thouless, 1977). Later it has been shown that

$$G \propto \exp(-L/L_{loc})$$

where $L_{loc} \sim \ell$ in the first case and $(bp_F/\hbar)^2 \ell$ for the second one (exponential localization). It seems that such a law is also the case for a metallic *film* (rigorous proof is absent).

Very simple-minded explanation - over-barrier reflection + *interference* of incoming and reflected waves. Because of the interference the condition $T = 0$ is crucial (no phase breaking). This explanation is good for one-dimensional case.

A little bit more scientific discussion. Consider interference corrections to the conductance at $T \rightarrow 0$. In this case one has to replace

$$\tau_\varphi \rightarrow L^2/D, \quad L_\varphi \rightarrow L.$$

One can conclude that in 3d case the relative correction is $\sim (\hbar/p_F \ell)^2 \ll 1$ (usually). However, at $d = 1, 2$ it *increases* with the size.

At what size $\Delta\sigma \approx \sigma$?

$$L_c = \begin{cases} \ell \exp(p_F^2 b \ell / \hbar^2), & d = 2 \\ \ell (p_F \ell / \hbar)^2, & d = 1 \end{cases} \quad (2.18)$$

We can conclude that in 3d case localization takes place at $p_F \ell \sim \hbar$, while in 1d and 2d case it takes place at any concentration of impurities.

Scaling hypothesis – According to the “gang of 4” (Abrahams, Anderson, Licciardello and Ramakrishnan)

$$G(qL) = f[q, G(L)]. \quad (2.19)$$

Assuming $q = 1 + \alpha$, $\alpha \ll 1$, we can iterate this equation in α :

$$\begin{aligned} G(L) &= f[1, G(L)]; \\ \alpha LG'(L) &= \alpha(\partial f / \partial q)_{q=1}. \end{aligned} \quad (2.20)$$

Denoting

$$G^{-1}(\partial f / \partial q)_{q=1} \equiv \beta(G)$$

we re-write the scaling assumption through the Gell-Mann & Low function, $\beta(G)$:

$$\partial \ln G / \partial \ln L = \beta(G).$$

At very large G we can expect that the usual theory is valid:

$$G = \sigma \begin{cases} S_{\perp}^2 / L = L, & d = 3; \\ bL_{\perp} / L = b, & d = 2; \\ b^2 / L, & d = 1 \end{cases} \quad (2.21)$$

Thus, in the zero approximation,

$$\beta(G) = d - 2.$$

Then we can use the weak localization approximation to find next corrections. One can get

$$\beta(G) \approx d - 2 - \alpha_d G_0 / G, \quad (2.22)$$

where

$$G_0 = e^2 / (\pi^2 \hbar), \quad \alpha_d \sim 1.$$

Indeed, for $d = 3$

$$\ln G = \ln[(\sigma + \Delta\sigma)L] \approx \ln \sigma L + (\Delta\sigma)/\sigma \approx \ln[(\sigma + \Delta\sigma|_{L=\infty})L] + \hbar^2 / (p_F^2 \ell L).$$

Thus

$$\beta(G) = \frac{\partial \ln G}{\partial \ln L} = 1 - \frac{\hbar^2}{p_F^2 \ell L} = 1 - \frac{\hbar^2 \sigma}{p_F^2 \ell G} = 1 - \alpha_3 \frac{G_0}{G}.$$

At small G one can suggest exponential localization:

$$G \sim G_0 \exp(-L/L_c) \quad \rightarrow \quad \beta(G) \sim \ln(G/G_0).$$

Thus we have the scenario shown in Fig. 2.14. Believing in such a scenario we get localization for 1d case ($\beta(G) < 0$ - conductance *increases* with the length). At $d = 3$ we have a fixed point at G_c , which is *unstable* (β changes sign). Under the simplest assumptions

$$\begin{aligned} \beta(G) &\approx \gamma(\ln G - \ln G_c) \approx \gamma(G/G_c - 1), \\ G &= G^{(0)} \quad \text{at } L = L_0 \text{ (initial condition)} \end{aligned}$$

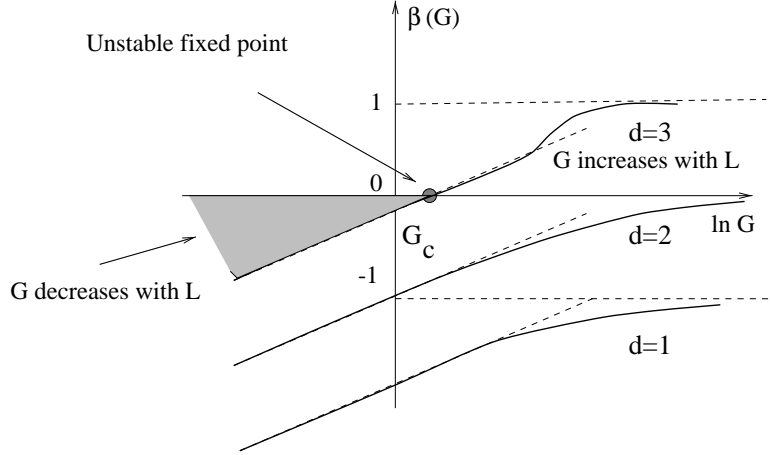


Figure 2.14: Flow curves for Anderson localization.

($G^{(0)}$ is close to G_c) we obtain

$$\frac{G}{G_c} \approx \left(\frac{G^{(0)}}{G_c} \right)^{(L/L_0)^\gamma}. \quad (2.23)$$

From Eq. (2.23) one can find important dependencies near the critical point. It is natural to choose $L_0 \approx \ell$ and to suggest that at this size

$$\sigma_0 \sim e^2 p_F^2 \ell / \hbar^3 \quad \rightarrow \quad G^{(0)} = \sigma_0 L_0 = (e^2 / \hbar) (p_F \ell / \hbar)^2.$$

Now we can assume that we can control some parameter (say, impurity content, x , which effects the mean free path ℓ), and that $G^{(0)}$ is regular in this parameter. Then, at small x $G^{(0)} = G_c(1 + x)$. At $L_0 = \ell$ we obtain

$$G = G_c(1 + x)^{(L/\ell)^\gamma} \approx G_c \exp[x(L/\ell)^\gamma].$$

Of course, such an assumption valid only at

$$x \ll 1, \quad L/\ell \gg 1.$$

Thus at $x < 0$ we obtain exponential localization with the characteristic length

$$L_{\text{loc}} \sim \ell |x|^{-1/\gamma}, \quad x < 0.$$

However, at $x > 0$ G grows with L . In a spirit of the concept of phase transitions we can treat the quantity $L_c = \ell x^{-1/\gamma}$ as a *correlation length*. At the scales of the order of L_c the properties of conducting and insulator phases are similar. The above law can be valid only in the vicinity of $G^{(0)} \sim G_c$. Then we match the usual Ohm's law, $G = \sigma L$. Consequently, the conductivity can be estimated as

$$\sigma \sim \frac{G^{(0)}}{L_c} \sim \sigma_0 x^{1/\gamma}, \quad x > 0.$$

Thus, we predict *power law*. Experiments on the dielectric constant ($\kappa_0 \propto L_c^2$) support the value

$$\gamma = 0.6 \pm 0.1.$$

Don't forget that we discuss the case $T = 0$, and the conductance is supposed at zero frequency. The range of applicability is given by the inequalities

$$L_c \ll L_\varphi, \quad L_\omega = \sqrt{D/\omega}.$$

This is almost impossible to meet these conditions, so usually people *extrapolate* experimental curves to $T = 0$, $\omega = 0$. It is a very subtle point because, as it was shown, the conductance at $L \leq L_\varphi$ is not a *self-averaging* quantity with respect to an ensemble of samples. More precise, the fluctuations between the samples *cannot be described* by the Gaussian law, their distribution being much wider. Then,

- have all these scaling assumptions any sense?
- Why they reasonably agree with the experiment?

The answer is positive because both the scaling predictions and the experiment are valid as an extrapolation from the region $L \geq L_\varphi$.

As the temperature grows, fluctuations decrease and the conductance tends to the Ohm's law.

Role of $e - e$ interaction

Nobody can consider both disorder and interaction acting together. To get some understanding of the role of $e - e$ interaction let us consider a *clean* metallic conductor. Assume that under some external perturbation (say, pressure) the band overlap changes. In this way we control the Fermi level (number of electrons and holes). One can consider them as free ones as their kinetic energy $p_F^2/2m^*$ exceeds the potential energy $e^2/(\kappa_0 \bar{r})$. In this way we come to the condition

$$\frac{e^2}{\kappa_0 \hbar v} \leq 1.$$

Otherwise electrons and hole form complexes - Wannier-Mott excitons. This state is insulating because excitons are neutral. This is only one of possible scenario. In general, the problem is a front end of modern condensed matter physics.

Chapter 3

Impurity band for lightly doped semiconductors.

The material is called lightly doped if there is only a small overlap between electronic states belonging to different impurities, $Na^3 \ll 1$. Here N is the impurity concentration while a is the localization length. In lightly doped materials conductivity vanished *exponentially* as $T \rightarrow 0$.

Let us start with a material with only one type of impurities, say n -type. At $T = 0$ each of donors must have an electron, the missing electron represents an *elementary excitation*. An excitation can be localized at any donor, the energies being slightly different. The reasons are

- If one removes an electron, the remaining positive charges polarize the neutral donor located in vicinity. That contributes to donor ionization energy, the contribution being dependent on the configuration of neutral neighbors.
- There is a quantum overlap between the donors being excited.

The first mechanism is usually more important for realistic semiconductors having compensating impurities.

Now let us assume that there are also acceptors which capture electrons from the donors and are fully occupied $T = 0$. Thus we have neutral donors, negatively charged acceptors and an equal number of positively charged donors. These randomly distributed charges create a fluctuating random potential which localizes the electronic states.

Let us count the energy ϵ_i of i -th donor from the energy of an isolated donor, $-E_0$. We have

$$\epsilon_i = \frac{e^2}{\kappa} \left[\sum_l^{\text{acc}} \frac{1}{|\mathbf{r}_i - \mathbf{r}_l|} - \sum_{k \neq i}^{\text{don}} \frac{1 - n_k}{|\mathbf{r}_i - \mathbf{r}_k|} \right].$$

The occupation numbers have to be determined to minimize the free energy (at $T = 0$ – electrostatic energy).

A typical dependence of the Fermi level μ on the degree of compensation, $K = N_A/N_D$, is shown in Fig. 3. Below we discuss limiting cases of small and large K .

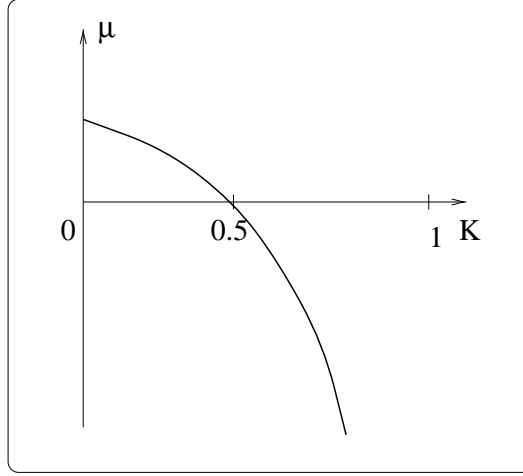


Figure 3.1: Position of the Fermi level μ as a function of the degree of compensation, K .

3.1 Low degree of compensation

At $K \ll 1$ most of donors keep their electrons, and only a small number is empty and charged positive. Each acceptor carries a negative charge, the concentration of charged donors is equal to N_A .

A positive charged hole will be located as close as possible to a negative acceptor. Thus, it occupies the donor the most closed to an acceptor. In a first approximation, each acceptor can be regarded as immersed in an infinite sea of donors. Transporting a hole from a donor situated near a charged acceptor to infinity requires work of order $e^2/\kappa r_D$, where r_D is the average distance between the donors.

To get a quantitative theory one has to add more input. Namely, there are some acceptor without holes, and some acceptors having 2 holes. Indeed, some acceptors do not have a close donors (Poisson distribution), so they bind holes very weakly. So holes prefer to find another acceptor with more close donors even at the expense to become a second hole. To illustrate the situation let us consider a configuration shown in Fig. 3.1,a. The work necessary to remove a hole equals to

$$e^2/\kappa r - e^2/2\kappa r = e^2/2\kappa r,$$

it becomes large at small r .

It is curious that an acceptor cannot bind more than 2 holes. Consider a configuration shown in Fig. 3.1,b. The energy of attraction to the acceptor equals $\sqrt{3}e^2/\kappa l$ while the repulsion energy from two other holes is $2e^2/\kappa l$. We end at the following situation. There are 3 configurations – 0-complexes (negative) where there is no ionized donor near a particular acceptor, 1-complexes (neutral), and 2-complexes (positive). So the neutrality condition which fixes the Fermi level is actually

$$N_0(\mu) = N_2(\mu).$$

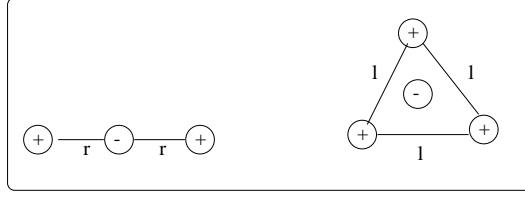


Figure 3.2: Donor-acceptor configurations

For 0-complex, there must be *no donors* within the sphere of radius $r_\mu = e^2/\kappa\mu$ from a fixed acceptor, the probability being $\exp(-4\pi N_D r_\mu^3/3)$. Thus,

$$N_0(\mu) = N_A \exp\left(-\frac{4\pi}{3} \frac{e^6 N_D}{\kappa^3 \mu^3}\right).$$

It is much more difficult to find number of 2-complexes. Let us estimate it *from above*, $N_2^>(\mu)$, as a concentration of pairs of donors whose additional energies ϵ_1 and ϵ_2 exceed μ when *both donors* are ionized. This estimate is larger because

- it could be another close donor which initiates a 1-complex,
- it can be more than one pair of donors near a given acceptor.

Let us put the origin of coordinates at acceptor site and suppose that a donor is located at \mathbf{r}_1 . The average number of donors in the element $d\mathbf{r}_2$ is equal to $N_D d\mathbf{r}_2$. Thus, the number of pairs with $r_2 \geq r_1$ is

$$N_D \int_{r_2 > r_1} d\mathbf{r}_2 \Theta[\epsilon_1(\mathbf{r}_1, \mathbf{r}_2) - \mu] \Theta[\epsilon_2(\mathbf{r}_1, \mathbf{r}_2) - \mu].$$

Here

$$\epsilon_{1,2} = \frac{e^2}{\kappa} \left[\frac{1}{|\mathbf{r}_{1,2}|} - \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right].$$

To get the total concentration of pairs one has to multiply it by N_D and integrate over \mathbf{r}_1 , and finally multiply by N_A . As a result,

$$N_2^>(\mu) = N_A N_D^2 \int d\mathbf{r}_1 \int_{r_2 > r_1} d\mathbf{r}_2 \Theta[\epsilon_1(\mathbf{r}_1, \mathbf{r}_2) - \mu] \Theta[\epsilon_2(\mathbf{r}_1, \mathbf{r}_2) - \mu].$$

This estimate is very close to exact result. Integration yields

$$N_2^>(\mu) = 7.14 \cdot 10^{-4} N_A (4\pi N_D r_\mu^3/3)^2.$$

Using this estimate and solving the neutrality equation one obtains

$$\mu = 0.61\epsilon_D, \quad \epsilon_D = e^2/\kappa r_D = (e^2/\kappa)(4\pi N_D/3)^{1/3}.$$

Now let us discuss the shape of the peak. The majority of donors are far from few acceptors, so there must be a sharp peak at $-E_0$. The tails of $g(\epsilon)$ above the Fermi level falls with the characteristic energy ϵ_D . Consequently, near the Fermi level

$$g(\epsilon) \propto N_A/\epsilon_D.$$

Long-range potential

Above we did not take into account electrostatic interaction between the complexes. It is small indeed because $N_0 \ll N_D$. However, the interaction leads to an additional dispersion of the peak.

An average fluctuations of charge produce an average potential of order $e^2 N_0^{1/3} / \kappa$. We shall show that *long-range* fluctuations are much more important. Let us introduce fluctuations of the complex concentrations for 0- and 2-complexes,

$$\xi_i(\mathbf{r}) = N_i(\mathbf{r}) - \langle N_i(\mathbf{r}) \rangle .$$

We consider the fluctuations as uncorrelated,

$$\begin{aligned} \langle \xi_0(\mathbf{r}) \xi_0(\mathbf{r}') \rangle &= \langle \xi_2(\mathbf{r}) \xi_2(\mathbf{r}') \rangle = N_0 \delta(\mathbf{r} - \mathbf{r}') , \\ \langle \xi_0(\mathbf{r}) \xi_2(\mathbf{r}') \rangle &= 0 . \end{aligned}$$

Consequently, the charge fluctuations are

$$\langle \rho(\mathbf{r}) \rho(\mathbf{r}') \rangle = 2e^2 N_0 \delta(\mathbf{r} - \mathbf{r}') . \quad (3.1)$$

Let us now consider a sphere of radius R where there is $\sim N_0 R^3$ complexes. The typical charge in the sphere is $e^2 (N_0 R^3)^{1/2}$, the resultant potential being $e^2 (N_0 R^3)^{1/2} / R$. It diverges as $R \rightarrow \infty$, and one has to allow for *screening*. We will check later that the screening potential varies little over a distance between the complexes. Therefore, one can employ the self-consistent field approximation.

As a result, we end at the Poisson equation

$$\Delta \phi = -\frac{4\pi}{\kappa} \rho(\mathbf{r}) - \frac{4\pi e}{\kappa} [N_2(\mu + e\phi) - N_0(\mu + e\phi)] .$$

Here $\rho(\mathbf{r})$ is a Gaussian random function whose correlator is given by (3.1). This approach is consistent at

$$|e\phi(\mathbf{r})| \ll \mu$$

when only small number of complexes are responsible for screening. At the same time, the majority of complexes contribute to charge fluctuations and are uncorrelated. In such a way we obtain the equation

$$\Delta \phi = -\frac{4\pi \rho}{\kappa} + \frac{\phi}{r_0^2} \quad (3.2)$$

where the screening length r_0 is given by the expression

$$\frac{1}{r_0^2} = \frac{4\pi e}{\kappa} \frac{d}{d\phi} [N_2(\mu + e\phi) - N_0(\mu + e\phi)]_{\phi=0} .$$

Substituting the expressions for the complex concentrations we obtain

$$r_0 = 0.58 N_A^{-1/2} N_D^{1/6} = 0.58 N_D^{-1/3} K^{-1/2} .$$

Then one can calculate the potential distribution from Eq. (3.2) as

$$e\phi(\mathbf{r}) == \int \rho(\mathbf{r}') K(\mathbf{r} - \mathbf{r}') d\mathbf{r}', \quad (3.3)$$

where

$$K(\mathbf{r}) = \frac{e}{\kappa r} e^{-r/r_0}. \quad (3.4)$$

We have

$$\begin{aligned} \langle e\phi(\mathbf{r}) \rangle^2 &= \int d\mathbf{r}_1 \int d\mathbf{r}_2 \langle \rho(\mathbf{r}_1) \rho(\mathbf{r}_2) \rangle K(\mathbf{r} - \mathbf{r}_1) K(\mathbf{r} - \mathbf{r}_2) \\ &= 2e^2 N_0 \int K^2(\mathbf{r}) = \frac{e^2}{\kappa r_0} \sqrt{2\pi N_0 r_0^3}. \end{aligned} \quad (3.5)$$

In a standard way, one can determine the distribution function near the Fermi level. It is Gaussian,

$$F(e\phi) = \frac{1}{\gamma\sqrt{\pi}} \exp[-(e\phi)^2/\gamma^2], \quad \gamma = 2e^2 \langle \phi(\mathbf{r}) \rangle^2 = 0.26\epsilon_D K^{1/4}.$$

All the above results are applicable only at very small K . Indeed, to get a Gaussian fluctuations one has to assume $r_0 \gg N_0^{-1/3}$ or

$$4\pi N_0 r_0^3 / 3 \gg 1.$$

The condition can be rewritten as $K^{-1/2} \ll 0.05$.

A typical sketch of the impurity band for a weakly compensated semiconductor is shown in Fig. 3.3

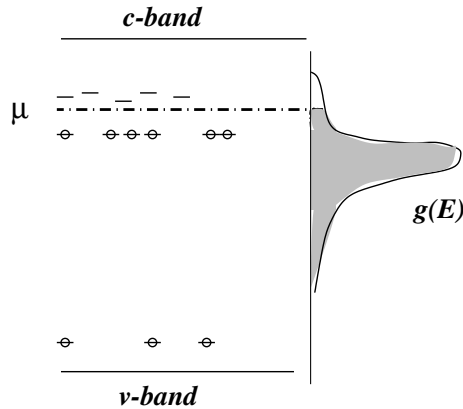


Figure 3.3: Energy diagram for a weakly compensated semiconductor

3.2 High degrees of compensation

Now we turn to the case

$$1 - K \ll 1.$$

In this case, the concentration of occupied donors,

$$n = N_D - N_A \ll N_D.$$

Thus *all the electrons* are able to find proper donors whose potential is lowered by the potential of charges neighborhood. Thus the Fermi level is situated well *below* $-E_0$, see Fig. 3.4

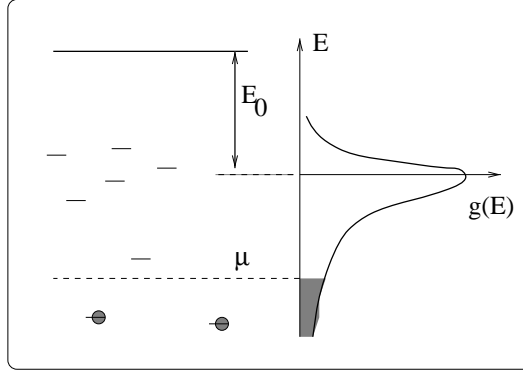


Figure 3.4: Energy diagram of highly compensated semiconductor. Solid line shows the conduction band edge, occupied states are shaded.

To understand the structure of the DOS tail consider a certain donor having at small ($r \ll N_D^{-1/3}$) another donor which is positively charged. Its influence at $r \gg a$ is just $\epsilon = -e^2/\kappa r$. It is assumed that the second donor is empty, the binding energy being $\sim E_0$. Thus a pair can contain only 1 electron.

To find $g(\epsilon)$ let us calculate the concentration of pairs having the energy in the interval $(\epsilon, \epsilon + d\epsilon)$. The probability to find another donor is $4\pi N_D r^2 (dr/d\epsilon) d\epsilon$, where $r(\epsilon) = e^2/\kappa\epsilon$. One has to multiply this quantity by N_D and divide by 2 (pairs!) to get

$$g(\epsilon) = \frac{3}{2} \frac{\epsilon_D^3}{\epsilon^4} N_D, \quad \epsilon_D = \frac{e^2}{\kappa} (4\pi N_D/3)^{1/3}.$$

Now the Fermi level (at $T = 0$) is easily found from the conservation law

$$\int_{-\infty}^{\mu} g(\epsilon) d\epsilon = n.$$

As a result, we get

$$\mu = -\frac{\epsilon_D}{2^{1/3}(1-K)^{1/3}}.$$

This is purely classical result. It is valid only at $r_\mu \gg a$.

Long-range fluctuations

At $1 - K \ll 1$ it is a large and important effect because screening is weak. To obtain an estimates we can repeat the discussion for $K \ll 1$ and replace the donor concentration by the total concentration, N_t , of donors and acceptors. In this way we get for a typical potential energy of an electron,

$$\gamma(R) = \frac{e^2}{\kappa R} (N_t R^3)^{1/2}.$$

It also diverges at large R .

How the screening takes place. The excess fluctuating density is

$$\Delta N = (N_t R^3)^{1/2} / R^3.$$

They can be neutralized at $\Delta N = n$ leading to the expression for the screening length,

$$n = \frac{(N_t r_s^3)^{1/2}}{r_s^3} \rightarrow r_s = \frac{2^{2/3} N_t^{-1/3}}{(1 - K)^{2/3}}.$$

The random potential amplitude is

$$\gamma(r_s) = \frac{e^2 N_t^{2/3}}{\kappa n^{1/3}}.$$

It increases with decreasing n .

Here we disregarded the states with several extra electrons which are important very deep in the gap.

3.3 Specific features of the two-dimensional case

In high-mobility heterostructures the two-dimensional electron gas is separated from the doped region by a relatively wide undoped region, *spacer*, with thickness s . The fluctuations of the charged donors create a random potential $F_b(\mathbf{r})$ in the 2DEG plane. The reason to introduce a spacer is that the random potential becomes smooth and transport cross-section a small factor $2p_F s / \hbar$ due to low-angle scattering.

To calculate this potential it is practical to assume that the spatial distribution of charged donors in the doped plane is random and not correlated,

$$\langle C(\mathbf{r}_1) \rangle = 0, \quad \langle C(\mathbf{r}_1) C(\mathbf{r}_2) \rangle = C \delta(\mathbf{r}_1 - \mathbf{r}_2). \quad (3.6)$$

For the field $F_b(\mathbf{r})$ we obtain

$$F_b(\mathbf{r}) = \frac{e^2}{\kappa} \int \frac{C(\mathbf{r}') d^2 r'}{\sqrt{|\mathbf{r} - \mathbf{r}'|^2 + s^2}} = 2\pi \frac{e^2}{\kappa} \int \frac{d^2 q}{q} C(\mathbf{q}) e^{i\mathbf{q} \cdot \mathbf{r} - qs}. \quad (3.7)$$

Here

$$C(\mathbf{q}) = \int \frac{d^2r}{(2\pi)^2} C(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} \rightarrow \langle C(\mathbf{q}_1)C(\mathbf{q}_2) \rangle = C(2\pi)^2 \delta(\mathbf{q}_1 + \mathbf{q}_2). \quad (3.8)$$

As a result, we obtain

$$\langle F_b^2(\mathbf{r}) \rangle = 2\pi C \int_{q_{\min}}^{\infty} \frac{dq}{q} e^{-2qs}. \quad (3.9)$$

Here $q_{\min} \sim 1/L$ where L is the linear dimension of the sample. Denoting

$$W \equiv \frac{e^2}{\kappa} \sqrt{2\pi C} \quad (3.10)$$

we have

$$\sqrt{\langle F_b^2 \rangle} \approx W \sqrt{\ln(L/2s)}. \quad (3.11)$$

For GaAs heterostructure with $C = 1.0 \times 10^{11} \text{ cm}^{-2}$ we have $W = 9.2 \text{ meV}$. That means very large fluctuations of bare potential in a macroscopic sample.

Electron screening removes this divergence. In the linear approximation it can be accounted for by replacing q by $q + q_s$ in the denominator of the integrand of Eq. (3.9)

$$\langle F_b^2(\mathbf{r}) \rangle = 2\pi C \int_{q_{\min}}^{\infty} \frac{dq}{q + q_s} e^{-2qs}. \quad (3.12)$$

where

$$q_s = 2\pi \frac{e^2}{\kappa} \frac{dn}{d\mu}, \quad (3.13)$$

where n is the electron density while μ is the chemical potential of electrons. Following previous derivation we obtain

$$\sqrt{\langle F_b^2 \rangle} = \frac{W}{2(q_s s)}, \quad \text{for } q_s s \gg 1. \quad (3.14)$$

We obtain that the linear screening dramatically reduces the potential since $q_s = 2/a_B$ where a_B is the Bohr radius. However, this estimate does not hold if the relative change in concentration is not small relative to the average concentration. Then the fluctuations in the long-range potential are of the order W .

The main idea of nonlinear screening is that entire plane screens each Fourier harmonic of the impurity charge distribution independently. The excess average square density of impurities in a square of area $R \times R$ is $\sqrt{CR^2}/R^2 = \sqrt{C}/R$. If the electron density n is much greater, $n \gg \sqrt{C}/R$, then the harmonic with the scale R will be screened linearly. However, the harmonics with $R \ll \sqrt{C}/n$ are not screened effectively, and 2DEG density becomes strongly inhomogeneous at the scale R . As a result, we arrive at the picture that harmonics in the impurity distribution with wavelengths $R \gg R_c = \sqrt{C}/n$ are screened linearly, while those with $R \ll R_c$ are screened very poorly.

For large electron density, $R_c \ll s$, and all significant harmonics in the bare potential are screened linearly. If the electron density is lowered, the nonlinear screening length

becomes larger, and eventually, for small enough electron density R_c becomes of the same order as s . At small densities, the fluctuations include all harmonics between s and R_c , and quantitative results can be obtained only numerically. We come to a conclusion that at some electron density a *percolation* between metallic droplets disappears, and a *metal-to-insulator transition* takes place. An of the critical electron density is

$$n_c\beta = \frac{\sqrt{C}}{s}, \quad \beta \approx 0.1. \quad (3.15)$$

An extremely interesting problem appears in an external magnetic field. We will not discuss it here.

Chapter 4

DC hopping conductance in lightly doped semiconductors: General description

4.1 Basic experimental facts

At large temperatures semiconductors possess an *intrinsic* electrical conductivity due to thermal activation of carriers across the energy gap. In this case the electron, n , and hole, p , concentrations are equal and exponentially temperature dependent, be

$$n = p = \frac{(2\pi kT \sqrt{m_e m_h})^{3/2}}{4\pi^3 \hbar^3} e^{-E_g/2kT}. \quad (4.1)$$

Due to the activation energy $E_g/2$ the concentration rapidly decreases with the temperature. At low temperatures it becomes less than the concentration contributed by impurities. In this region the conduction is called *extrinsic* since it depends on the impurities. A schematic plot of temperature dependence of the conductance is shown in Fig. 4.1 The region *A* corresponds to intrinsic conduction, while the regions *B* – *D* correspond to extrinsic conductance. If the impurities are shallow, then there exists the region *B* which is called the saturation range. In this range, all the impurities are ionized and hence the carrier concentration is temperature independent. Consequently, the temperature dependence of resistivity is determined entirely by the carrier *mobility*. For example, a decrease in resistivity can be explained by a weaker phonon scattering.

Further decrease in the temperature leads to a gradual freeze-out of the electrons which are recaptured by donors (for *n*-type material).

Experimentally, the temperature dependence of the resistivity in the region *C* often shows only one activation energy, ϵ_1 ,

$$\rho(T) = \rho e^{\epsilon_1/kT}, \quad (4.2)$$

where ϵ_1 is close to the ionization energy E_0 of an isolated donor.

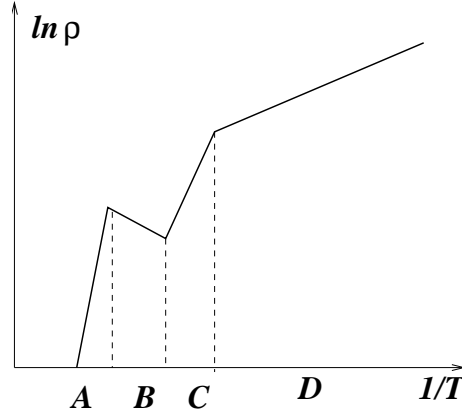


Figure 4.1: Schematic temperature dependence of the resistivity of a lightly doped semiconductor. *A* - region of intrinsic conductance, *B* saturation region of impurity concentration, *C* - freeze-out region, *D* - hopping conduction region.

At lower temperatures, hopping between the impurities without excursion to the conduction band becomes most important (region D). The experimental results for *p*-type neutron-doped Ge is shown in Fig. 4.2. Each curve corresponds to a different acceptor (Ga) concentration. As a result, an interpolation formula for the regions C and D has the form

$$\rho^{-1}(T) = \rho_1^{-1} e^{-\epsilon_1/kT} + \rho_3^{-1} e^{-\epsilon_3/kT}. \quad (4.3)$$

The 1st item corresponds to the band conductivity, and it is almost independent of the acceptor concentration. The second one corresponds to hopping conduction (it will clear later why we gave the subscript 3 for it). We see that that hopping conduction has a noticeable activation energy ϵ_3 , though it is small compared to ϵ_1 . It arises due to the dispersion of impurity levels. In hopping over donors, an electron emits and absorbs phonons that results in an exponential temperature dependence.

When impurity concentration increases, it first somehow enhances that activation energy. However, a further increase in the concentration increases an overlap between the wave functions of neighboring centers. At $N_A = 10^{17} \text{ cm}^{-3}$ the activation energy ϵ_3 vanishes, and the conductance crosses over to a metallic one.

A characteristic feature of hopping which is evident from Fig. 4.2 is an extremely strong dependence of ρ_3 on impurity concentration. It can be represented as

$$\rho_3 = \rho_{03} e^{f(N_D)} \quad (4.4)$$

where ρ_{03} and $f(N_D)$ are power-law functions of the impurity concentration. The exponential dependence is rather clear - it is just due to change of the wave-function overlap integrals.

What is also important is that both ρ_3 and ϵ_3 are strong functions of the *degree of compensation*. The dependence $\epsilon_3(K)$ is non-monotonous, it decreases sharply at first, then has a minimum at $K \approx 0.4$, and then grows rapidly at $K \rightarrow 1$.

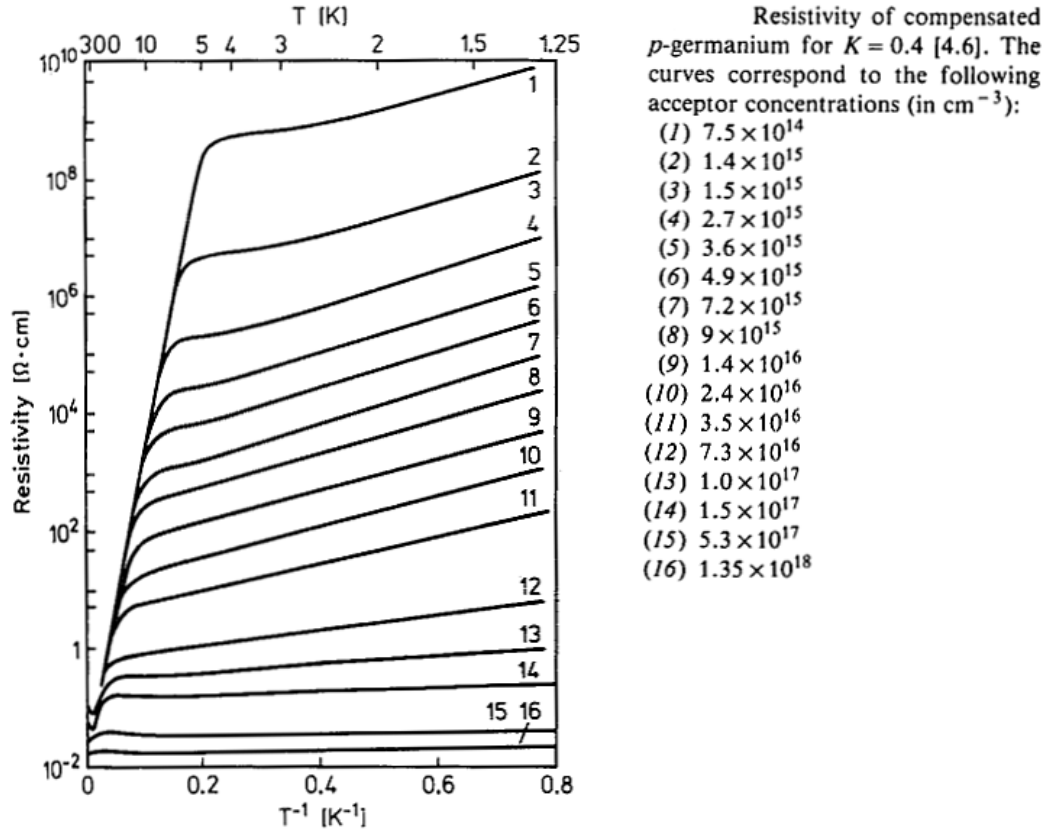


Figure 4.2:

In addition to the band and hopping mechanisms of conduction, semiconductors with low compensation $K \leq 0.2$ display another activated mechanism which manifests itself near the Mott transition. The mechanism contributes one more term of the form

$$\rho_2^{-1} e^{-e_2/kT}.$$

We will not discuss this feature in detail.

4.2 The Abrahams-Miller resistor network model

4.2.1 Derivation

To calculate the hopping conductance we shall use the so-called resistor network model introduced by *Miller* and *Abrahams*. First, starting from electron wave functions we shall calculate the hopping probability between the centers i and j with the absorption and emission of a phonon. Then we shall calculate the number of $i \rightarrow j$ transition per unit time. In the absence of an electric field there is a detailed balance, and this number is exact equal to the number of reverse transitions. However, in the presence of the of a weak

electric field an imbalance appears, and a current appears proportional to the voltage drop between the centers. Consequently, we can introduce an effective resistance R_{ij} of a given transition, and thus the whole problem is reduced to calculating of effective resistance of a network of random resistors. This way will be traced in detail below.

Consider two donors having the coordinates \mathbf{r}_i and \mathbf{r}_j and sharing 1 electron. For simplicity, let us assume that the distance between the donors is large, so the overlap will be weak. In the absence of the overlap the wave function $\Psi_i \equiv \Psi(\mathbf{r} - \mathbf{r}_i)$ satisfies the Schrödinger equation with the potential $U = -e^2/\kappa|\mathbf{r} - \mathbf{r}_i|$. Within the framework of linear combinations of atomic orbitals (LCAO) method, the system is represented by symmetric and asymmetric combinations of atomic wave functions,

$$\Psi_{1,2} = \frac{\Psi_i \pm \Psi_j}{\sqrt{2} \sqrt{1 \pm \int d\mathbf{r} \Psi_i^* \Psi_j}} \quad (4.5)$$

The corresponding energies are obtained from the Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 - \frac{e^2}{\kappa|\mathbf{r} - \mathbf{r}_i|} - \frac{e^2}{\kappa|\mathbf{r} - \mathbf{r}_j|} \quad (4.6)$$

are

$$E_{2,1} = -E_0 - \frac{e^2}{\kappa r_{ij}} \pm I_{ij}, \quad (4.7)$$

where E_0 is the energy of an isolated donor, while I_{ij} is the *energy overlap integral* is

$$I_{ij} = \int \Psi_i^* \Psi_j \frac{e^2}{\kappa|\mathbf{r} - \mathbf{r}_j|} d\mathbf{r} - \int \Psi_i^* \Psi_j d\mathbf{r}' \int \frac{e^2 |\Psi_i|^2}{\kappa|\mathbf{r} - \mathbf{r}_j|} d\mathbf{r}. \quad (4.8)$$

Now let us consider a simple case where the donor state is connected to a single extreme in the Brillouin zone center. Then we can split integral into a sum over crystal cells and use the fact that the envelope function is almost constant over a cell. The same can be done with potential energy which contains $|\mathbf{r} - \mathbf{r}_j| \gg a_0$. The cell integrals $|u_{n,0}|^2$ give unity. In this way we come to the expression

$$I_{ij} = \int F_i(\mathbf{r}) F_j(\mathbf{r}) \frac{e^2}{\kappa|\mathbf{r} - \mathbf{r}_j|} d\mathbf{r} - \int F_i(\mathbf{r}') F_j(\mathbf{r}') d\mathbf{r}' \int \frac{e^2 |F_i|^2}{\kappa|\mathbf{r} - \mathbf{r}_j|} d\mathbf{r}. \quad (4.9)$$

For a hydrogen-like function

$$F(r) = \frac{1}{\sqrt{\pi a^3}} e^{-r/a}$$

we get

$$I_{ij} = \frac{2}{3} \left(\frac{e^2}{\kappa a} \right) \left(\frac{r_{ij}}{a} \right) \exp \left(-\frac{r_{ij}}{a} \right). \quad (4.10)$$

Note that there is a more accurate calculation which does not assume the overlap integral to be small. It differs from Eq. (4.10) by replacing $2/3 \rightarrow 2/e$. The general feature of the energy overlap integral is that it can be expressed as

$$I_{ij} = I_0 \exp \left(-\frac{r_{ij}}{a} \right) \quad (4.11)$$

where I_0 is of the order of the effective Bohr energy and only weakly dependent on r_{ij} .

In important fact is that the electron interacts also with the other charged impurities surrounding the center by some potential $W(\mathbf{r})$. As a result, for the majority of donors,

$$\Delta_i^j \gtrsim I_{ij}, \quad \text{where} \quad \Delta_i^j \equiv W(\mathbf{r}_j) - W(\mathbf{r}_i). \quad (4.12)$$

Consequently, the potential W should be included into the Hamiltonian. The simplest situation case is when $\Delta_i^j \gg I_{ij}$. Then for the two lowest states we can show that

$$\begin{aligned} \Psi'_i &= \Psi_i + \frac{I_{ij}}{\Delta_i^j} \Psi_j, \\ \Psi'_j &= \Psi_j - \frac{I_{ij}}{\Delta_i^j} \Psi_i. \end{aligned} \quad (4.13)$$

Despite of the the fact that an admixture of the “foreign” state is small, it leads to the charge transfer on distance r_{ij} .

Now let us consider phonon-induced transitions and assume that only a single branch of acoustic phonons is present. The transition probability $i \rightarrow j$ with *absorption* of one phonon is

$$\gamma_{ij} = \frac{2\pi}{\hbar} \frac{\mathcal{V}_0}{(2\pi)^3} \int |M_{\mathbf{q}}|^2 \delta(\hbar s q - \Delta_i^j) d\mathbf{q}, \quad (4.14)$$

where \mathcal{V}_0 is the volume of the crystal, s is the speed of sound, \mathbf{q} is the phonon wave vector, while

$$M_{\mathbf{q}} = iD \sqrt{\frac{\hbar q N_{\mathbf{q}}}{2\mathcal{V}_0 \rho}} \int \Psi'_j e^{i\mathbf{q} \cdot \mathbf{r}} \Psi'_i d\mathbf{r} \quad (4.15)$$

is the matrix element of electron-phonon interaction. Here D is the deformation potential constant, $N_{\mathbf{q}}$ is the Planck function, ρ is the crystal density. Substituting the functions Ψ' and coming to the envelope functions we rewrite the previous expression in the form

$$\begin{aligned} M_{\mathbf{q}} &= iD \sqrt{\frac{\hbar q N_{\mathbf{q}}}{2\mathcal{V}_0 \rho}} \left\{ \frac{I_{ij}}{\Delta_i^j} \left[\int F_j^2 e^{i\mathbf{q} \cdot \mathbf{r}} d\mathbf{r} - \int F_i^2 e^{i\mathbf{q} \cdot \mathbf{r}} d\mathbf{r} \right] \right. \\ &\quad \left. + \left[1 - \left(\frac{I_{ij}}{\Delta_i^j} \right)^2 \right] \int F_i F_j e^{i\mathbf{q} \cdot \mathbf{r}} d\mathbf{r} \right\}, \end{aligned} \quad (4.16)$$

To estimate the integral let us evaluate two dimensionless parameters,

$$q r_{ij} = \Delta_i^j r_{ij} / \hbar s \quad \text{and} \quad q a = \Delta_i^j a / \hbar s.$$

Since $\Delta_i^j \approx e^2 N_D^{1/3} \kappa^{-1}$ and $R_{ij} \approx N_D^{-1/3}$ we see that $q r_{ij}$ is usually large (20-30) while $q a$ is of the order 1. The characteristic length of variation of $F_i F_j$ across the line connecting two centers is of the order $\sqrt{r_{ij} a}$. Thus the last integral in Eq. (4.16) is strongly suppressed due to oscillatory factor, and it can be neglected. The first two integrals can be combined in the form

$$(e^{i\mathbf{q} \cdot \mathbf{r}_{ij}} - 1) \int F_i^2 e^{i\mathbf{q} \cdot \mathbf{r}} d\mathbf{r} = \frac{(e^{i\mathbf{q} \cdot \mathbf{r}_{ij}} - 1)}{(1 + q^2 a^2 / 4)^2}.$$

The last equality is valid for hydrogen-like wave functions. Finally we obtain

$$|M_{\mathbf{q}}|^2 = \frac{\hbar q D^2}{2\rho \mathcal{V}_0 s} \left(\frac{I_{ij}}{\Delta_i^j} \right)^2 \frac{N_{\mathbf{q}}(1 - \cos \mathbf{q} \cdot \mathbf{r}_{ij})}{(1 + q^2 a^2/4)^4}. \quad (4.17)$$

Substituting this expression into (4.14) and using the fact that $qr_{ij} \gg 1$ we get the final expression

$$\gamma_{ij} = \gamma_{ij}^0 e^{-2r_{ij}/a} N(\Delta_i^j) \quad (4.18)$$

where

$$\begin{aligned} \gamma_{ij}^0 &= \frac{D^2 \Delta_i^j}{\pi \rho s^5 \hbar^4} \left(\frac{2e^2}{3\kappa a} \right)^2 \left(\frac{r_{ij}}{a} \right)^2 \frac{1}{[1 + (\Delta_i^j a / 2\hbar s)^2]^4}, \\ N(\Delta_i^j) &= [\exp(\Delta_i^j / kT) - 1]^{-1}. \end{aligned} \quad (4.19)$$

Let $n = (0, 1)$ are the occupancies of i th donor which fluctuate in time. The transition $i \rightarrow j$ is possible if $n_i = 1$, $n_j = 0$. Therefore the average number of electrons which make transitions is

$$\Gamma_{ij} = \langle \gamma_{ij} n_i (1 - n_j) \rangle. \quad (4.20)$$

The average should be taken with respect to time. Here we make a *fundamental assumption* that the occupancies can be replaced by their *average numbers*, $\langle n_i \rangle \equiv f_i$, while the energies can be replaced by self-consistent energies

$$\varepsilon_i = \sum_l^{\text{acc}} \frac{e^2}{\kappa |\mathbf{r}_i - \mathbf{r}_l|} - \sum_{k \neq i}^{\text{don}} \frac{e^2 (1 - f_k)}{\kappa |\mathbf{r}_i - \mathbf{r}_k|}. \quad (4.21)$$

In this approximation we can put $\Delta_i^j = \varepsilon_j - \varepsilon_i \equiv \varepsilon_{ij}$ to obtain,

$$\Gamma_{ij} = \gamma_{ij}^0 e^{-2r_{ij}/a} N(\varepsilon_j - \varepsilon_i) f_i (1 - f_j). \quad (4.22)$$

For the reverse process we get

$$\Gamma_{ji} = \gamma_{ij}^0 e^{-2r_{ij}/a} [N(\varepsilon_j - \varepsilon_i) + 1] f_j (1 - f_i). \quad (4.23)$$

The current between the centers j and i can be written as

$$J_{ij} = -e(\Gamma_{ij} - \Gamma_{ji}). \quad (4.24)$$

In the absence of an electric field

$$f_i = f_i^0 = \left[\frac{1}{2} \exp \left(\frac{\varepsilon_0 - \mu}{kT} \right) + 1 \right]^{-1} \quad (4.25)$$

where ε_0 is the average energy at the site at $\mathbf{E} = 0$, while factor $1/2$ enters because there are 2 possible spin states at the site i . In fact the energies ε_0 and the functions f_i^0 are made self-consistent by Eq. (4.21).

It can be easily seen that a detailed balance is present, and there is no current. The electric field, first, redistributes the electrons between the donors, $f_i^0 \rightarrow f_i^0 + \delta f_i$. We describe these changes by small additions $\delta\mu_i$,

$$f_i(\mathbf{E}) = \left[\frac{1}{2} \exp \left(\frac{\varepsilon_0 - \delta\mu_i - \mu}{kT} \right) + 1 \right]^{-1} \quad (4.26)$$

Secondly, the field affects the donor level energies,

$$\varepsilon_i = \varepsilon_i^0 + \delta\varepsilon_i, \quad \delta\varepsilon_i = e\mathbf{E} \cdot \mathbf{r}_i + \frac{e^2}{\kappa} \sum_{k \neq i}^{\text{don}} \frac{\delta f_k}{|\mathbf{r}_i - \mathbf{r}_k|}. \quad (4.27)$$

Would the circuit be broken, a new equilibrium with $\delta\mu_i + \delta\varepsilon_i = 0$ will result. However in the case of conductance, $\delta\mu_i + \delta\varepsilon_i \neq 0$. Assuming all the correction to be small (*linear in \mathbf{E} effects*) we can expand all the energies up to the first order in $\delta\varepsilon_i$ and $\delta\mu_i$. After such manipulation we obtain

$$J_{ij} = \frac{e\Gamma_{ij}^0}{kT} [\delta\mu_i + \delta\varepsilon_i - (\delta\mu_j + \delta\varepsilon_j)] \quad (4.28)$$

The latter equation can be expressed as the Ohm's law

$$J_{ij} = R_{ij}^{-1}(U_i - U_j) \quad (4.29)$$

where

$$R_{ij} = kT/e^2\Gamma_{ij}^0, \quad \Gamma_{ij}^0 = \gamma_{ij}^0 e^{-2r_{ij}/a} N(\varepsilon_j - \varepsilon_i) f_i^0(1 - f_j^0), \quad (4.30)$$

$$-eU_i = \delta\mu_i + \delta\varepsilon_i = e\mathbf{E} \cdot \mathbf{r}_i + \frac{e^2}{\kappa} \sum_{k \neq i}^{\text{don}} \frac{\delta f_k}{|\mathbf{r}_i - \mathbf{r}_k|}. \quad (4.31)$$

The quantity $-eU_i$ can be regarded as a local value of the electrochemical potential on the donor i counted from the chemical potential μ . Thus $U_i - U_j$ is the voltage drop on the transition $i \rightarrow j$ while R_{ij} is its resistance. We end at the problem of calculating of effective resistance of a random resistor network keeping the potentials at the electrodes constant.

4.2.2 Conventional treatment

Let us analyze the expressions for the resistances R_{ij} and assume the temperatures to be low,

$$kT \ll |\varepsilon_i^0 - \varepsilon_j^0|, |\varepsilon_i^0 - \mu|, |\varepsilon_j^0 - \mu|.$$

Then we can extract the functions strongly dependent on the distance r_{ij} and on the energies,

$$\Gamma_{ij}^0 \approx \gamma_{ij}^0 e^{-2r_{ij}/a} e^{-\varepsilon_{ij}/kT} \quad (4.32)$$

where

$$\epsilon_{ij} = \frac{1}{2} (|\varepsilon_i - \varepsilon_j| + |\varepsilon_i - \mu| + |\varepsilon_j - \mu|) . \quad (4.33)$$

Now it is convenient to introduce the quantity

$$\xi_{ij} = \frac{2r_{ij}}{a} + \frac{\epsilon_{ij}}{kT} \quad (4.34)$$

to express random resistors as

$$R_{ij} = R_{ij}^0 \exp(\xi_{ij}) , \quad R_{ij}^0 = kT/e^2\gamma_{ij}^0 . \quad (4.35)$$

The main feature of the problem is an exponentially wide range of resistances. It is clear that doth averaging of conductances and resistances will lead to wrong answers. A correct method of averaging is suggested by the mathematical theory of percolation.

Chapter 5

Percolation theory

The term “percolation” has been introduced in 1957 by *Broadbent* and *Hammersley* in connection with a new class of mathematical problems. We shall start with general description and then will map them to the hopping conductance.

5.1 Lattice problems

5.1.1 “Liquid flow” definition

Imagine an infinite lattice with bonds between the adjacent sites which permit the flow in both directions, and that each wet site will instantly wet the adjacent sites. The various problems follow from introduction random elements into the theory.

Bond problem

Assume that each bond in the lattice can be either blocked, or unblocked. Let the probability of any bond to be *unblocked* is “ x ”. The distribution of unblocked and blocked bonds remains constant in time.

If a random site is wet, then there are 2 possibilities. The initial site will wet either a finite or an infinite number of other sites. The key quantity is the *probability* of a random initial site to wet an *infinite* number of other sites. In the infinite lattice this probability depends only on x . Let us denote it as $P^b(x)$. The bond problem is illustrated in Fig. 5.1 (left panel), while the function $P^b(x)$ is plotted for different lattices in the right panel. Curves 1,2 and 4 correspond to 3D lattices, the rest to 2D ones. When x is small $P^b(x) \equiv 0$ since the blocked bonds prevent the liquid to spread far from the initial site. On the other hand, as x approaches 1, so does $P^b(x)$. A very important quantity is the *percolation threshold*, x_c , defined as an upper limit for the x -values for which $P^b(x) = 0$. When $x - x_c \ll 1$, one has

$$P^b(x) \sim (x - x_c)^\beta \tag{5.1}$$

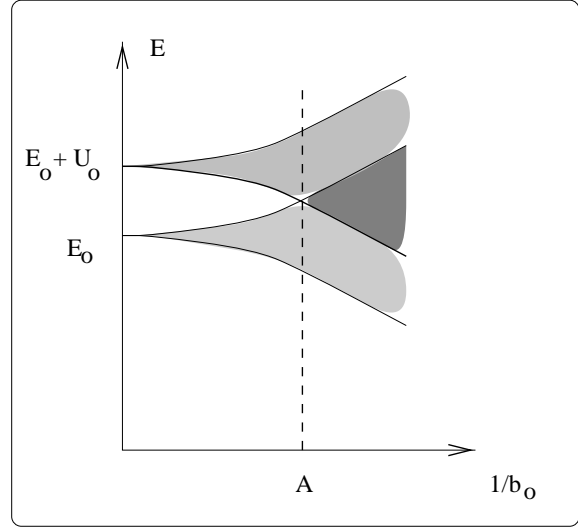
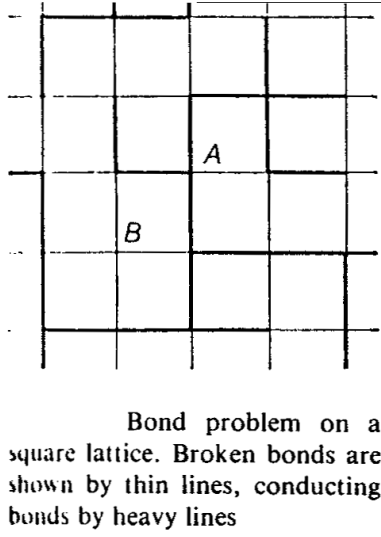


Figure 5.1:

where b is a *critical exponent*. The situation resembles type II phase transitions. Note that all the conclusions are valid for an *infinite* lattice.

One can also define the probability to wet N other sites, $P_N^b(x)$. $P_N^b(x)$ differs from 0 at $0 < x < 1$, but at $N \gg 1$ it is small for $x < x_c$. We can define

$$P^b(x) = \lim_{N \rightarrow \infty} P_N^b(x).$$

An example of the bond problem - an orchard which is planned in the form of a square lattice. It is known that a diseased tree contaminates another tree away at the distance r with the probability $f(r)$. It is required to find a minimum lattice period h_{\min} to avoid epidemic. The solution is obvious,

$$f(h_{\min}) = x_c. \quad (5.2)$$

Site problem

For the site problem, *sites* can be either blocked, or unlocked, while the bond are assumed to be ideal. Blocked sites permit no flow in any direction, they cannot be wet. Let x be the fraction of *unblocked* sites while $P^s(x)$ be a probability for a random site to wet infinite number of other sites.

The probability function $P^s(x)$ is plotted in the left panel of Fig. 5.2, while a realization of the site problem is shown in the right panel. One can see that there is a percolation threshold which depends in the lattice type.

5.1.2 Cluster statistics

Both the site and both problems are usually discussed in terms of *cluster statistics*, rather in terms of liquid flow. Let us apply this approach to the site problem. Imagine that a

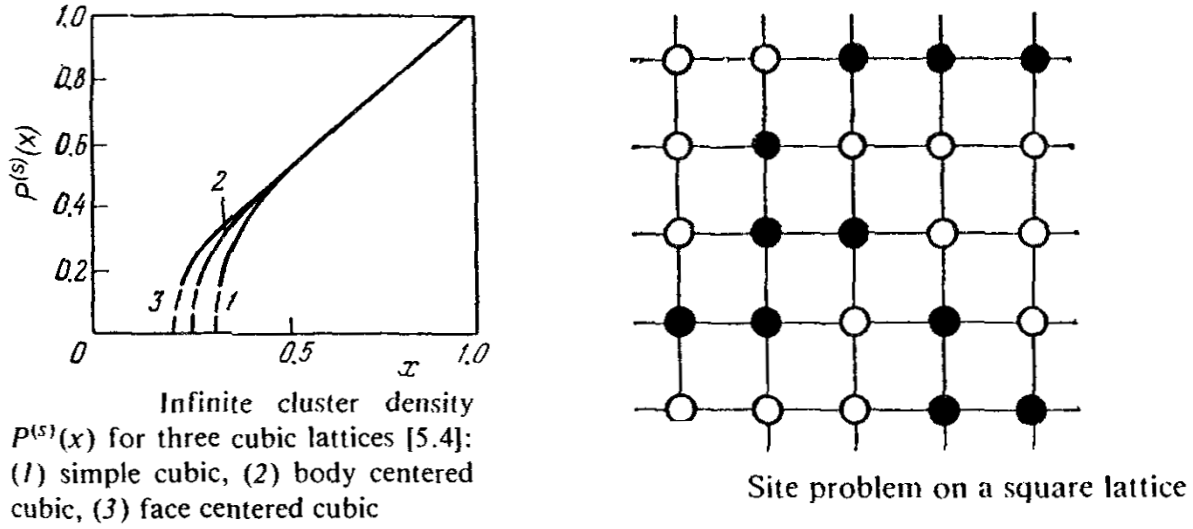


Figure 5.2:

fraction x of the sites is painted black, while the rest sites are painted white. Any two adjacent sites are connected if both are black. A chain of such connected clusters becomes a *cluster*. For example in the right panel of Fig. 5.2 there are 3 clusters - one consists of 5 black sites while two consist of 3 black sites.

When x is small, then there are only small clusters. As we approach the percolation threshold, several clusters may come together, and the mean cluster size *increases*. At $x \rightarrow x_c$ an *infinite* black cluster is born. It resembles a random network which permeates the entire lattice, while smaller since finite clusters exist in the “pores”, see Fig. 5.3. Actually, $P^s(x)$ has a meaning of the *density of the infinite cluster*. Similarly, a proper

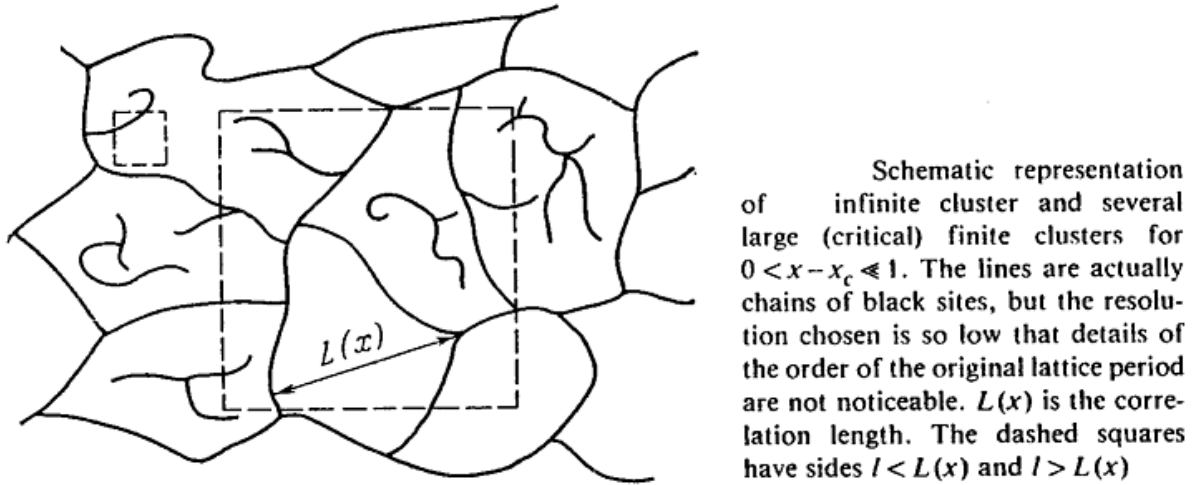


Figure 5.3:

definition for the bond problem can be formulated.

It is assume that *no more than one* infinite cluster can exist in a lattice. Than brings us to an important application of the site problem - a crystalline solution of ferromagnetic atoms (with a fraction “ x ”) in a non-magnetic host. Assuming nearest-neighbor ferromagnetic interaction we arrive at the ferromagnetic phase transition at zero temperature as a percolation threshold for the site problem. Consequently,

- the transition takes place at $x = x_c$;
- the saturation magnetization at $T = 0$ for $x > x_c$ can be expressed through the density of the infinite cluster,

$$M|_{T=0} = \mu_0 P^s(x) \quad (5.3)$$

where μ_0 is the atomic magnetic moment.

Another important application is the aforementioned Anderson transition. The appearance of the extended states is similar to formation of an infinite black cluster.

Now let us introduce relevant quantities for the cluster statistics. Let n_s be the number of clusters containing s sites, per site of an infinite lattice. Then the fraction of sites belonging to an infinite cluster is

$$P(x) = x - \sum_s s n_s. \quad (5.4)$$

The second important quantity is the *mean cluster size*,

$$S(x) = \frac{\sum_s s^2 n_s}{\sum_s s n_s}. \quad (5.5)$$

According to numerical estimates, $S(x)$ becomes infinite at x_c ,

$$S(x) \sim |x - x_c|^{-\gamma}, \quad (|x - x_c| \ll 1), \quad (5.6)$$

where γ is a critical exponent. Note that Eq. 5.5 calculates a weighted average which is divergent. A simple average would be regular because

$$\lim_{x \rightarrow x_c} \sum_s s n_s = x_c.$$

The majority of sites remain in clusters with $s \sim 1$ even at $x = x_c$.

According to numerical estimates, n_s near the percolation threshold behaves as follows. There is a *critical number* s_c of the sites in the cluster, which grows at $x \rightarrow x_c + 0$ and at $x \rightarrow x_c - 0$,

$$s_c(x) \sim |x - x_c|^{-\Delta}, \quad \Delta > 0. \quad (5.7)$$

If $s \ll s_c$ and s increases the quantity n_s decreases according to some power law, while at $s \gg s_c$ it decreases exponentially. Consequently, the numerator of Eq. (5.5) is determined

by the clusters with $s \approx s_c$, which we will refer to as *critical clusters*. The divergence of $S(x)$ means that the number of sites in critical cluster increases.

The third important quantity in the percolation theory is the *correlation function*. Let us define a function $g(\mathbf{r}_i, \mathbf{r}_j)$ as equal to 1 if both sites are black and belong to the same cluster, and 0 otherwise. Now we can introduce the pair correlation function as

$$G(r, x) = G(\mathbf{r}_i - \mathbf{r}_j, x) = \langle g(\mathbf{r}_i, \mathbf{r}_j) \rangle. \quad (5.8)$$

Naturally, $G(r, x)$ does to zero at $r \rightarrow \infty$ because n_s decreases with s . At the distances smaller than the *average size of the critical cluster* which we will denote as $L(x)$, the correlation is governed by the clusters with $s \ll s_c(x)$. Consequently, $G(r, x)$ decreases as a power law of r . If $r \gg L(x)$, the exponentially rare clusters with $s \gg s_c(x)$ are dominant, and $G(r, x)$ decreases exponentially. The length to the critical clusters, $L(x)$, is called the *correlation length*. Since the number of sites in the critical clusters grow at $x \rightarrow x_c \pm 0$, the correlation length must grow if one approaches the percolation threshold,

$$L(x) \sim |x - x_c|^{-\nu}, \quad (|x - x_c| \ll 1). \quad (5.9)$$

Here ν is the critical exponent of the correlation length. It is usually thought that the critical exponents γ , Δ and ν are the same whether $x > x_c$ or $x < x_c$, and that there exists a sort of cluster-pore duality (see Fig 5.3). Thus, at $x < x_c$ the quantity $L(x)$ is a typical radius of clusters, while at $x > x_c$ it is a typical radius of the pores. Near x_c the correlation length introduces a typical distance (“period”) between the nodes of the network.

5.1.3 Percolation through a finite lattice

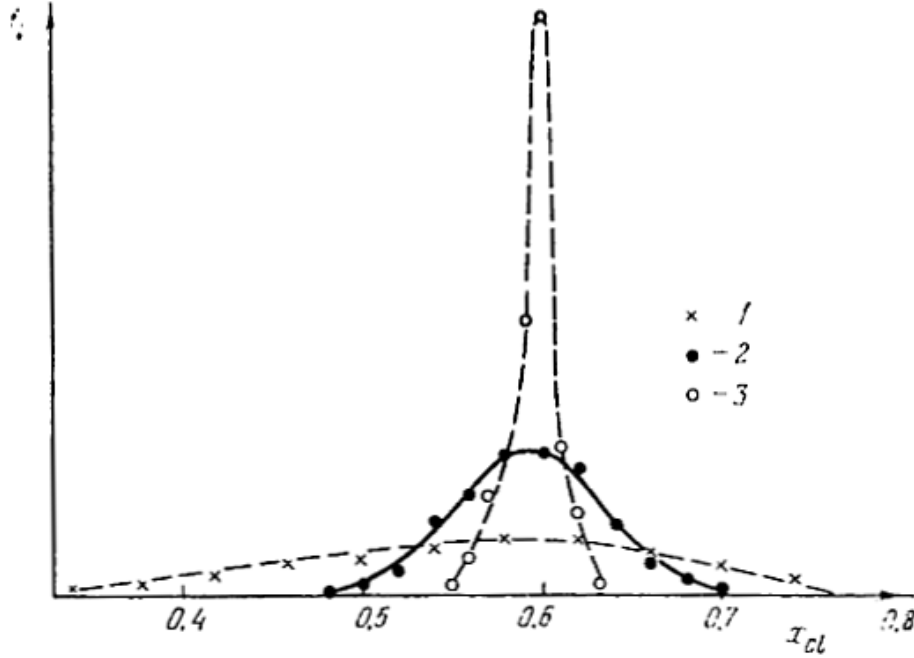
Consider a finite lattice containing ℓ sites per side. Let all the sites be initially white. We will randomly paint some sites black, slowly increasing the fraction x . At some $x = x_{c\ell}$ the black cluster will connect opposite sides. We will consider $x_{c\ell}$ as a percolation threshold for a finite lattice. If we permit the procedure, the arrangement of the black sites, as well as $x_{c\ell}$ will be different. Therefore, $x_{c\ell}$ is the random quantity, and we can define the true percolation threshold as

$$x_c = \lim_{\ell \rightarrow \infty} \langle x_{c\ell} \rangle. \quad (5.10)$$

Another variation of the same theme is to define x as the *probability* for a site to be black rather than as the fraction of the black sites. In a finite systems, these two definitions lead to different values of $x_{c\ell}$ and $\langle x_{c\ell} \rangle$ due to fluctuations. We can say that the 1st definition corresponds to a micro-canonical ensemble while the 2nd one - to the canonical ensemble.

Figure 5.4 shows the distribution function f_ℓ of $x_{c\ell}$ in a canonical ensemble for the site problem on a square lattice. This figure was obtained by Monte Carlo simulations. The distributions are nearly Gaussian in shape. The larger the side the lower the dispersion $W_\ell^2 = \langle (x_{c\ell} - \langle x_{c\ell} \rangle)^2 \rangle$. It was shown that W_ℓ obeys the power law

$$W_\ell = B\ell^{-1/\nu}, \quad (5.11)$$



The distribution functions of values of the percolation threshold for the site problem on $l \times l$ square lattices [5.13]. The l values in units of the lattice constant are: (1) 8, (2) 32, (3) 128

Figure 5.4:

while the average threshold obeys the law

$$\langle x_{cl} \rangle = x_c + A\ell^{-1/\lambda}. \quad (5.12)$$

In three dimensions, $\nu \approx 0.9$ and $\lambda \approx 1$. It can be shown that the exponent ν is the same as for the correlation length.

As an example of percolation between the opposite faces for the 3D bond problem let us consider electrical conductivity of a large cube in which the unblocked bonds are represented by standard resistors. The blocked bonds are assumed to have infinite resistance, Fig. 5.5. If the cube is large, then the conductivity differs from zero only at $x > x_c(b)$, while at $0 < x - x_c \ll 1$ one has

$$\sigma(x) \sim (x - x_c)^t, \quad (5.13)$$

where t is yet another critical exponent. It is clear that to find σ it is not enough to know $P^b(x)$ since also a topological structure of the infinite cluster is important.

5.1.4 Summary

We have presented 3 definitions of the percolation problem: liquid flow, cluster statistics, and percolation through a finite lattice. It can be shown that critical phenomena are the same in all three formulations. The typical questions to be answered:

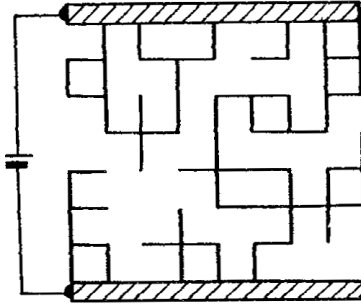


Figure 5.5: A random resistor network between two electrodes.

1. What is the percolation threshold x_c ?
2. What are the critical exponents β, γ, ν and t ?
3. How do dimensionality and lattice geometry affect the above mentioned quantities?
4. What is the topology of finite and infinite clusters?

The values of the percolation thresholds are summarized in Table 5.1 In the Table 5.2 the most important critical exponents are presented. The majority of values in the tables were found by numerical simulations, however some results are obtained analytically. In rare cases, experimental methods were used.

The modern point of view is that the *critical exponents* depend only on the dimensionality of the problem, and in this way they are *universal*.

5.2 Continuous problems

Suppose a random continuous function $V(\mathbf{r})$ is defined in the entire space by its correlations, such that $\langle V(\mathbf{r}) \rangle = 0$. Then let us paint black the regions where $V(\mathbf{r}) < V$ and leave other regions white. As V changes from $-\infty$ to ∞ the volume of black region changes from zero to all space, see Fig. 5.6. As V increases, the black regions grow, and at some V_c they merge into an “ocean”. The quantity V_c is called *the percolation threshold*. Such a “liquid-flow” definition can be easily reformulated into the “cluster statistics”, or “finite-volume percolation” ones.

One can easily map the continuous problem to the lattice one. Assume a lattice with so fine grid that the potential does not change. Then fix V and consider all the sites in the black regions to be black while the other sites - to be white. The the fraction of the black sites is

$$x = \theta(V) \equiv \int_{-\infty}^V F(V') dV' \quad (5.14)$$

Percolation threshold x_c for various lattices. Exact values are underlined. All other values were obtained by numerical methods				
Lattice	$x_c(b)$		$x_c(s)$	
Honeycomb	<u>0.6527</u>	[5.15,16]	$0.700 \pm .010$	[5.20]
	0.640 ± 0.018	[5.17-19]	0.688 ± 0.017	[5.17-19]
			0.697 ± 0.001	[5.21,22]
			0.698 ± 0.003	[5.23]
Square	<u>0.5000</u>	[5.15,16]	0.590 ± 0.010	[5.20]
	0.493 ± 0.013	[5.17-19]	0.581 ± 0.015	[5.17-19]
	0.498 ± 0.017	[5.21,22,24]	0.591 ± 0.001	[5.21,22]
			0.593 ± 0.02	[5.23]
			0.594 ± 0.03	[5.13]
			0.5898 ± 0.0008	[5.25]
Triangular	<u>0.3473</u>	[5.15,16]	<u>0.5000</u>	[5.15,16]
	0.341 ± 0.011	[5.17-19]	0.493 ± 0.018	[5.17-19]
	0.349 ± 0.010	[5.21,22]	0.500 ± 0.001	[5.21,22]
Diamond	0.388 ± 0.005	[5.20]	0.425 ± 0.015	[5.20]
	0.390 ± 0.011	[5.17-19]	0.436 ± 0.012	[5.21,22]
Simple cubic	0.247 ± 0.005	[5.20]	0.307 ± 0.010	[5.20]
	0.254 ± 0.013	[5.17-19]	0.325 ± 0.023	[5.17-19]
			0.320 ± 0.001	[5.21,22]
			0.312 ± 0.002	[5.4]
			0.312 ± 0.003	[5.26]
			0.312 ± 0.001	[5.27]
Body centered cubic	0.178 ± 0.005	[5.20]	0.243 ± 0.010	[5.20]
			0.254 ± 0.001	[5.21,22]
			0.248 ± 0.003	[5.4]
			0.248 ± 0.003	[5.26]
Face centered cubic	0.119 ± 0.002	[5.20]	0.195 ± 0.005	[5.20]
	0.125 ± 0.005	[5.17-19]	0.199 ± 0.008	[5.17-19]
	0.1190 ± 0.0005	[5.28]	0.208 ± 0.001	[5.21,22]
			0.200 ± 0.002	[5.4]
			0.198 ± 0.003	[5.29]

Table 5.1: From the book [1]

Exponent	$d = 2$		$d = 3$	
β	0.14 ± 0.03	[5.30]	0.35 ± 0.05	[5.4]
	0.138 ± 0.007	[5.31]	0.39 ± 0.02	[5.27]
	0.148 ± 0.004	[5.32]	0.42 ± 0.06	[5.29]
			0.47 ± 0.02	[5.33]
γ	1.85 ± 0.20	[5.32]		
	2.38 ± 0.03	[5.20]	1.69 ± 0.05	[5.20]
	2.38 ± 0.02	[5.28]	1.70 ± 0.11	[5.28]
	2.43 ± 0.03	[5.23]	1.80 ± 0.05	[5.27]
	2.34 ± 0.30	[5.34]		
ν	1.34 ± 0.02	[5.28]	0.94 ± 0.05	[5.13,26]
	1.33 ± 0.04	[5.13]	0.82 ± 0.05	[5.28]
	1.36 ± 0.04	[5.26]	0.83 ± 0.13	[5.35]
	≤ 1.5	[5.25]		
t			2	[5.41]
	1.38 ± 0.12	[5.36]	1.725 ± 0.005	[5.42]
	1.1 ± 0.1	[5.27]	1.6 ± 0.1	[5.27]
	1.15 ± 0.15	[5.37]	1.75 ± 0.10	[5.38]
	$1.1 - 1.25$	[5.38]	1.6 ± 0.1	[5.43]
	1.31 ± 0.04	[5.39]		
	1.26 ± 0.03	[5.40]		

Table 5.2: From the book [1]

where $\theta(V)$ is the space fraction of the black sites, while $F(V)$ is the distribution function of the quantity V . The percolation threshold then can be defined by equating the quantity

$$\theta_c = \int_{-\infty}^{V_c} F(V) dV \quad (5.15)$$

to the percolation threshold x_c of the site problem. Physically it means that the distributions of black and white sites are now *correlated* according to the function $V(\mathbf{r})$. This proposition is very important for practical calculation of θ_c .

Along with the above mentioned “percolation through black”, one can define a “percolation through white” with some threshold V'_c . In the two-dimensional case $V_c = V'_c = 0$, and $\theta_c = 0.5$. The situation in 3 dimensions is more complicated. Only an inequality $V_c \leq V'_c$ does exist, so the following 3 situations are possible:

1. $V < V_c$ and there is percolation through white, but not through black;
2. $V_c \leq V < V'_c$ and there is percolation through both black and white;
3. $V \geq V'_c$ and there is percolation through black, but not through white.

If the potential is statistically symmetric, then

$$V_c \leq 0, \quad V'_c \geq 0, \quad \text{and} \quad \theta_c \leq 0.5. \quad (5.16)$$

The quantity θ_c has been extensively studied for various distribution functions of the random potential. A typical value of θ_c for Gaussian potentials is $\theta_c = 0.17$.

5.3 Random site problems

The most important problems for hopping conductance are the so-called *random-site problems*. Random sites are the sites chaotic distributed in space, the average number N of sites per unit volume being known. If one introduced some function $\xi_{ij}(\mathbf{r}_{ij})$, the two sites are considered bonded if, for some number ξ ,

$$\xi_{ij} \leq \xi. \quad (5.17)$$

Eq. (5.17) is referred to as the *bonding criterion*. If two sites are bonded either directly, or via some other sites, then they form a cluster. It is required to find the percolation threshold ξ_c , i. e. the lowest bound of the parameter ξ which still permits an infinite cluster.

The simplest example of the bonding criterion is

$$r_{ij} \leq r, \quad (5.18)$$

which is met if the state j is located within the sphere of the radius r centered at the site i . Consequently, the problem has a very simple geometric interpretation: spheres of radius

r are constructed around all the sites, and we must find the lowest value $r = r_c$ which allows endless chain of sites in which each site lies inside the sphere about the preceding site, see Fig. 5.7. We will term the quantity r_c the *percolation radius*. Occasionally it is more convenient to construct spheres with radius $r/2$ and calculate the value that allows for an infinite cluster of overlapping spheres.

The percolation threshold r_c depends solely on the concentration N and is must be proportional to $N^{-1/3}$. In most cases people consider the threshold value

$$B_c = (4/3)\pi N r_c^3 \quad (5.19)$$

of the dimensionless parameter $B = (4/3)\pi N r^3$ which corresponds to an average number of bonds per site. This problem is often called the sphere problem, or the disk problem in 2D case (where $B_c = \pi \tilde{N} r_c^2$).

A physical implementation of this problem is a ferromagnetism of a dilute crystalline mixture of magnetic and non-magnetic atoms, where the exchange forced between magnetic atoms are modeled by a finite-radius interaction with the interaction radius r_{int} . If r_{int} is much greater than the lattice constant so that many sites fall into the sphere of interaction, then we arrive at the sphere problem. However, the problem is a reversed one: knowing r_{int} we have to find N . If the number of sites that fall into the interaction sphere is Z that for the lattices of any dimensionality we obviously have

$$\lim_{Z \rightarrow \infty} Z x_c = B_c. \quad (5.20)$$

The critical indices for the random site problems can be introduced in the same way as for the lattice model. Introducing n_s as the number of clusters of size s in a given volume we can rewrite scaling properties in a similar way in terms of $\xi - \xi_c$.

In some important cases the bonding criterion is fulfilled on a complex surface rather than on a sphere. Then it should be written in the form

$$\xi_{ij} \equiv \phi(r_{ij}) \leq \xi. \quad (5.21)$$

Here $\xi > 0$, and $\phi(r)$ is a monotonously increasing homogeneous function of x_{ij} , y_{ij} , and z_{ij} . The equation $\phi(r_{ij}) = \xi$ define a surface which replace the sphere in the constructions discussed above. In particular, we can define B_c as $N V_{\xi_c}$. It can be shown that the values B_c are the same for all the surfaces which can be obtained from each other via a linear coordinate transformation,

$$x'_k = \sum_{\ell} A_{k\ell} x_{\ell}$$

since such transformation does not change the topology of the system. Thus the threshold values B_c for ellipsoids and spheres are the same.

5.4 Electric conductivity of random networks of conducting bonds. Infinite cluster topology

As we have mentioned for the case of lattice models, the electric conductivity of a lattice of conducting and non-conducting bonds behaves as

$$\sigma(x) \sim (x - x_c)^t \quad (x \geq x_c). \quad (5.22)$$

The aim of this section is to discuss the critical exponent t . There are few ways to study this problem: an experimental (by punching holes in a graphite paper, cutting bonds in a metallic grid, etc), computer simulations, etc. It appears that the results both for site and bond problems are closed in value. It seems logical that these values will retain also for the continuum version. What we want to do now is to relate the conductivity to other properties of the infinite cluster, namely, with its density $P(x)$ and the “period” $L(x)$. It is clear that these quantities do not contain the full necessary information about $\sigma(x)$, and we need some assumptions to relate the critical exponents.

Dead ends

Consider a large cube containing ℓ units per side. If the infinite cluster resembles a cubic lattice then $1/3$ of the bonds form chains that connect the opposite edges of the cube, the number of bonds in these chains being $\approx (1/3)P(x)\ell^3$. Each chain has ℓ sites, therefore the number of parallel chains is $\approx (1/3)P(x)\ell^2$. Since the resistance of each bond is $\propto \ell$, the resistance R of the whole cube is $\propto \ell[P(x)\ell^2]^{-1}$. Consequently,

$$\sigma(x) \sim (R\ell)^{-1} \sim P(x) \sim (x - x_c)^\beta. \quad (5.23)$$

We arrive at the conclusion that $t = \beta$. However, we know that $t > \beta$, see Table 5.2. That leads to a conclusion that almost all the “mass” of the infinite cluster is concentrated in its *dead ends*, see Fig. 5.8. These chains contribute to $P(x)$, but not to $\sigma(x)$. The insufficiency of the infinite cluster can be also due to redundant

The nodes and links model

This is the most used model (Skal-Shklovskii-De Gennes, SSDG) assumes that there is a backbone network with mean size $L \sim (x - x_c)^{-\nu}$, and that this backbone carries dead ends. The part of the backbone is called *the link*. It is assumed that at least half of the links are not doubled. Each link can be twisted, so its length \tilde{L} is larger than L . In this way we can write at $x - x_c \ll 1$

$$\tilde{L} \sim (x - x_c)^{-\zeta}, \quad \text{where } \zeta \geq \nu. \quad (5.24)$$

A schematic representation of the SSDG model is shown in Fig. 5.9. To proceed let us first show that $\zeta = 1$. Let $x > x_c$ and cut each bond with the probability $(x - x_c)/x$. As a

result, the fraction of unbroken bonds will become $x[1 - (x - x_c)/x] = x_c$. Consequently, we arrive at the percolation threshold for the *original* lattice. Now notice that to cut a link we must cut one of its constituent bonds (we have assume that the links are not doubled!). Thus, the probability for a link to be broken is proportional to \tilde{L} and equal $\tilde{L}(x - x_c)/x$. We can assume the percolation threshold for a backbone to be equal to some y_c . Equating $\tilde{L}(x - x_c)/x = y_c$ we find that $\zeta = -1$. We note that in 2D case $\nu > 1$, and it means that SSDG model does not hold, and that duplication of links is crucially important. However, for $d = 3$ $\nu_3 \approx 0.9$, and $\tilde{L} \sim L^{1.1}$. Note that the twistness appears very weak comparing to the cases of the random walk, $\tilde{L} \sim L^2$ and of the self-avoiding random walk, $\tilde{L} \sim L^{5/3}$.

Now let us calculate $\sigma(x)$ If resistance of each bond is R_0 then the resistance of a cube with side L is $R_0(\tilde{L}/\ell_0)$ where ℓ_0 is the bond length. At the scales larger than L the infinite cluster is homogeneous. Therefore $\sigma(x)$ can be calculated as an effective conductivity of a cube with the side of the order of the correlation length. We get

$$\sigma(x) = R^{-1}L^{2-d} = R_0^{-1}(\tilde{L}/\ell_0)^{-1}L^{2-d} = R_0^{-1}\ell_0^{2-d}(x - x_c)^{1+\nu(d-2)}. \quad (5.25)$$

As a result,

$$t = 1 + \nu(d - 2). \quad (5.26)$$

For $d = 3$ we get $t_3 = 1 + \nu_3$. Substituting $\nu_3 = 0.9$ we obtain $t_3 = 1.9$ while the numerical value is 1.7. We conclude that SSDG model is not very accurate. However it is extensively used because it is simple and qualitatively correct. It has been extensively used in the theory of dilute ferromagnets, granular superconductors, conductivity and Hall effect in disordered systems. The model has been also extended to take in some way into account the bond doubling. It makes the model more accurate.

The most important extension (Pike and Stanley) is to discriminate between *red* and *blue* backbone bonds. Breaking a red bond breaks the link itself. It was further suggested that blue bonds from *blobs* connected by singly connected chains. The number of red bond between the nodes has also be scaled as $L' \sim (x - x_c)^{-1}$. A finite size of the blobs removes the limitation of the model for the cases when $\nu > 1$, in particular in 2D case. It seems that the combined **nodes-links-blobs model** describes properly the structure of the infinite cluster.

Scaling hypothesis and calculation of $\sigma(x)$

Another way to estimate the exponent t is to suggest that the large-scale structure of the network remains self-similar as the system approaches the percolation threshold. The second assumption to be made is that the resistance is determined only by a large-scale structure. The two assumptions allow one to relate t and ν in a simple way.

Let us denote $\tau = x - x_c$ and examine the systems at two values of τ , such that $\tau_2 < \tau_1$. Since $L \sim |\tau|^{-\nu}$, the dimensions of the system will increase by the factor $L(\tau_2)/L(\tau_1) = (\tau_1/\tau_2)^\nu$. We assume that we also scale the distance between the contacts and the potential difference by the same factor. We conclude that the current distribution in the bond will not be affected. Now consider an area element that is perpendicular to the average direction

of the current. The number of channel piercing the area will decrease as $L(\tau_1)/L(\tau_2) = (\tau_2/\tau_1)^{2\nu}$. Thus the current and the conductivity will decrease by the factor $(\tau_1/\tau_2)^{2\nu}$. As a result, in three dimensions $t = 2\nu$, while in two dimensions $t = \nu$. In this way we obtain the values $t_3 = 1.80$ and $t_2 = 1.33$ close to the experimental ones. The scaling approach has appeared extremely fruitful during last years and revealed fractal properties of the infinite cluster.

5.5 Percolation theory and electric conductivity of strongly inhomogeneous media

We have previously shown that calculation of hopping conductivity can be reduced to conductivity of a random network with an exponentially wide range of resistances. The similar approach is valid also for a continuum.

Suppose that we have a medium in which local conductivity fluctuates as

$$\sigma(\mathbf{r}) = \sigma e^{-\xi(\mathbf{r})}, \quad \text{and} \quad (5.27)$$

$$\langle [\xi(\mathbf{r}) - \langle \xi \rangle]^2 \rangle \gg 1. \quad (5.28)$$

For example, the source of fluctuation can be a band bending due to a random large potential, and $\xi(\mathbf{r}) = \varepsilon(\mathbf{r})/kT$. Here $\varepsilon(\mathbf{r})$ is the distance between the Fermi level and the bottom of the conduction band. A similar problem can be also formulated on a resistor lattice with $R = R_0 e^{\xi'}$ where the random quantity ξ' is randomly distributed between $-\xi_0$ and $\xi_0 \gg 1$.

To illustrate the method of conductivity calculation let us consider a cubic lattice. Let us first derive the exponential factor in the effective conductivity. Then we will discuss also a pre-exponential factor, and finally compare the result with the SSDG model.

Consider a simple cubic lattice and select a value ξ in the interval $(-\xi_0, \xi_0)$. Then let us replace all resistances with $\xi' > \xi$ by *infinite* resistances thus breaking the circuits. Let us call the so-obtained conductivity as $\sigma(\xi)$. The chosen ξ determines the probability of a random resistance not being broken:

$$x(\xi) = \int_{-\xi_0}^{\xi} F(\xi') d\xi', \quad (5.29)$$

where the distribution function F is defined as

$$F(\xi) = \begin{cases} 1/2\xi_0, & |\xi| \leq \xi_0 \\ 0, & |\xi| > \xi_0 \end{cases}. \quad (5.30)$$

From Eqs. (5.29),(5.30) we find

$$x(\xi) = \frac{\xi_0 + \xi}{2\xi_0}. \quad (5.31)$$

Now let us increase the value of ξ from $-\xi_0$ till some value ξ_c corresponding to the percolation threshold for the bond problem, $x_c(b)$, given by Eq. (5.31). If we that increase ξ

from ξ_c to $\xi_c + 1$, the presence of infinite cluster will cause a rapid increase in conductivity. This happens because the correlation length rapidly diminishes as

$$L(\xi) \sim \ell_0 [x(\xi) - x_c(\xi)]^{-\nu} \sim \ell_0 \xi_0^\nu (\xi - \xi_c)^{-\nu}. \quad (5.32)$$

Therefore, the number of parallel conducting chains in the infinite cluster network will also rapidly increase. On the other hand, the change in the individual resistance remains small if ξ is changed less than 1. This implies a rapid power-law increase in σ ,

$$\sigma(\xi) \sim (\xi - \xi_c)^b, \quad \text{where } b > 0. \quad (5.33)$$

Let the infinite cluster that is formed when ξ and ξ_c differ by about one be called the critical subnetwork. Its resistance is determined by the highest resistances because, by definition these resistances cannot be shunted by lower ones since that would allow for percolation at $\xi < \xi_c$. Consequently, for a critical subnetwork we have

$$\sigma(\xi_c + 1) \sim \sigma_0 e^{-\xi_c}. \quad (5.34)$$

Further increase of ξ will not raise σ significantly, even though the density of infinite cluster will rise. Indeed, if we continue to increase ξ we introduce exponentially large resistances. Thus, new chains will be shunted by the critical subnetwork. The expression

$$\sigma = \sigma_0^{-\xi_c} \quad (5.35)$$

is the foundation for the hopping conduction theory. According to the Table 5.2, $x_c(b) = 0.25$ for the simple cubic lattice. From (5.31) we find $\xi_c = -\xi_0/2$, and

$$\sigma = \sigma_0^{\xi_0/2} \quad (5.36)$$

This formula is valid only if $\xi_0 \gg 1$, when the concept of the critical subnetwork is valid. Numerical tests show that the model is practical at $\xi_0 \geq 9$.

Let us now discuss the pre-exponential factor σ_0 . From dimensional arguments,

$$\sigma_0 = c(\xi_0)(R_0 \ell_0)^{-1} \quad (5.37)$$

where $c(\xi_0)$ is a dimensionless factor which can contain only powers of ξ_0 . It can be shown that this power equals to $-\nu$ where ν is the correlation length exponent.

To obtain the power, let us suggest another derivation of the conductivity. Consider a cube with side ℓ cut from the infinite lattice. Let us break all the resistances and then connect them back in the ascending order until we reach the percolation between the opposite sides. Let the last (greatest) resistance be $\xi_{c\ell}$. Therefore,

$$R = R_0 e^{\xi_{c\ell}} \rightarrow \sigma = (R_0 \ell)^{-1} e^{-\xi_{c\ell}}.$$

The quantity $\xi_{c\ell}$ is actually random, however for a large system its fluctuations are small. As we have seen, the fluctuations behave as

$$\frac{\delta \xi_{c\ell}}{\xi_0} \sim \left(\frac{\ell_0}{\ell} \right)^{1/\nu}.$$

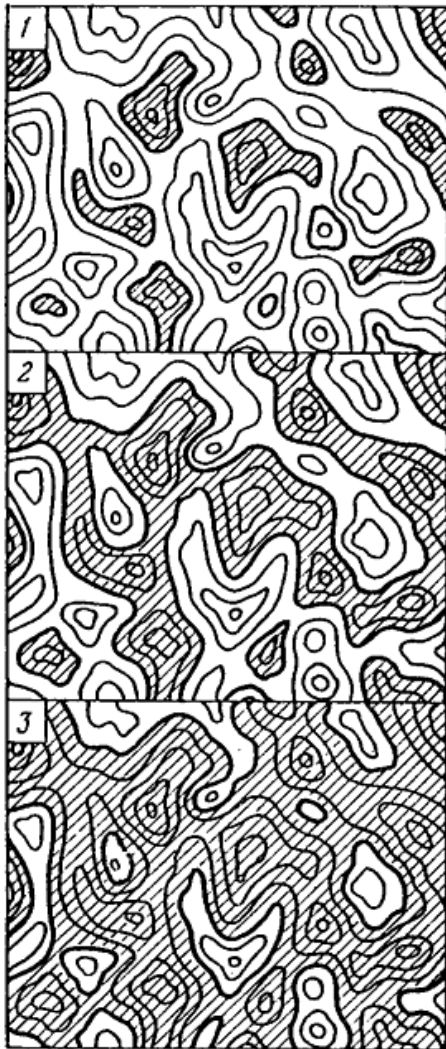
Now let us increase the cube size ℓ . At some size the conductivity becomes ℓ -independent. We can tell that it happens if the mean square fluctuation approaches unity, or at

$$\ell \approx \ell_0 \xi_0^\nu \equiv L_0.$$

L_0 is just the infinite cluster correlation length. For such a cube the difference between $\xi_{c\ell}$ and ξ_c is negligible. Thus,

$$\sigma(\xi_0) \sim (R_0 \ell_0)^{-1} \xi_0^{-\nu} e^{-\xi_c}.$$

A same expression can be easily derived from the node-link model of the infinite cluster.



Percolation in a two-dimensional potential [5.59]. The outlines correspond to surface $V(\mathbf{r}) = \text{const.}$ "Black" regions [$V(\mathbf{r}) < V$] are shaded. The three maps correspond to three different values of V ($V_1 < V_2 < V_3$)

Figure 5.6:

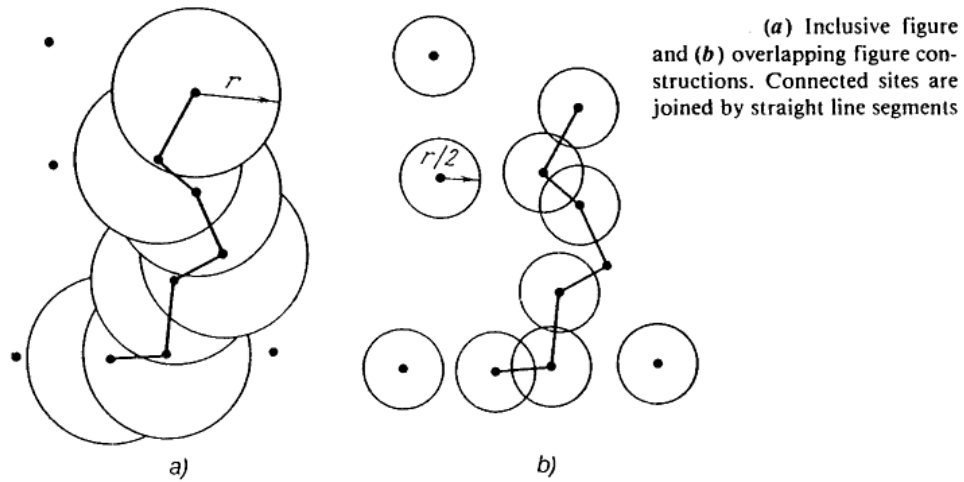


Figure 5.7:

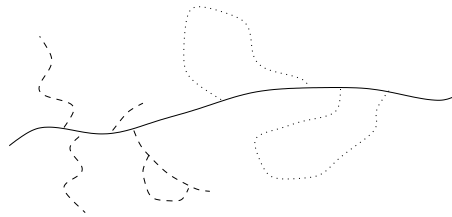


Figure 5.8: Dead ends (dashed) and redundant loops (dotted) on a conducting chain.

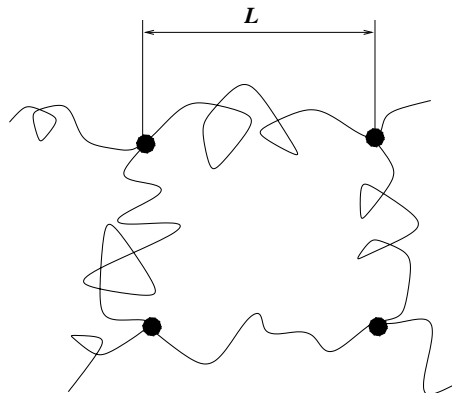


Figure 5.9: Schematic representation of the SSDG model.

Chapter 6

Hopping conductance

6.1 Dependence on impurity concentrations

In this section, we shall apply the percolation theory to study the dependence of the ρ_3 contribution on the impurity concentration.

6.2 The contribution ρ_3 as a function of impurity concentrations

Let us concentrate on the simplest case of isotropic impurity wave function. In fact, we have to calculate the resistance of the Abrahams-Miller network of random resistors,

$$R_{ij} = R_{ij}^0 e^{\xi_{ij}}, \quad (6.1)$$

where

$$\xi_{ij} = 2r_{ij}/a + \epsilon_{ij}/kT, \quad (6.2)$$

$$R_{ij}^0 = kT/e^2 \gamma_{ij}^0, \quad (6.3)$$

$$\epsilon_{ij} = (1/2) (|\epsilon_i - \epsilon_j| + |\epsilon_i - \mu| + |\epsilon_j - \mu|). \quad (6.4)$$

To calculate the resistance let us apply the percolation approach. First suppose that only the resistances with $\xi_{ij} < \xi$ are switched on. Let us gradually increase ξ until we reach the condition of percolation over conducting resistances. That obviously occurs at $\xi = \xi_c$. Increasing ξ from ξ_c to $\xi_c + 1$ results in formation of the critical subnetwork of resistances. This subnetwork *shunts* all the resistors for which $\xi \geq \xi_c + 1$. Consequently, the exponential part of the hopping conductivity to the parameters of the problem is given by

$$\rho = \rho_0 e^{\xi_c}. \quad (6.5)$$

This is the starting point of the theory of hopping conductivity.

Let us discuss the case of rather high temperatures when we can neglect the item ϵ_{ij}/kT in Eq. (6.4). The bonding criterion has now a very simple form, $r_{ij} \leq a\xi/2$, i. e. the connectivity of a pair depends only on their separation. In this way we return to the problems of spheres discussed in the previous chapter, see Fig. 5.7. Therefore ξ_c can be expressed in terms of the percolation radius r_c , i. e. the percolation threshold for an auxiliary problem with the bonding criterion $r_{ij} \leq r$ as

$$\xi_c = 2r_c/a. \quad (6.6)$$

According to numerical calculations,

$$B_c \equiv (4\pi/3)Nr_c^3 = 2.7 \pm 0.1, \quad r_c = (0.865 \pm 0.015)N^{-1/3}. \quad (6.7)$$

Here N is the concentration of the majority dopant. Substituting (6.7) into (6.6), we find from (6.5)

$$\rho_3 = \rho_{03} \exp \left[\alpha / (N^{1/3}a) \right], \quad \alpha = 1.73 \pm 0.03. \quad (6.8)$$

This result agrees with the experiment. The problem to calculate the activation energy is more delicate and the answer depend strongly on the degree of compensation.

The case of anisotropic wave function is different just geometrically because the bonding condition becomes anisotropic.

A very interesting changes occur in an external magnetic field. The presence of magnetic field squeezes the electron wave functions, and the overlap integrals between the components of the pair decrease. Furthermore, the wave functions are squeezed mostly in the direction perpendicular \mathbf{H} . As a result the resistivity exponentially increases in a strong magnetic field, the resistivity becoming anisotropic. Giant positive magnetoresistance is the hallmark of hopping conductance.

6.3 Activation energy

6.3.1 Low degree of compensation

As we have seen, the resistivity of a lightly doped semiconductor in the hopping regime has the form

$$\rho = \rho_3 e^{\epsilon_3/kT}. \quad (6.9)$$

Since we were interested only in ρ_3 the energy term ϵ_{ij}/kT has been ignored in the expression for ξ_{ij} . To take it into account let write ξ_{ij} in the form $\xi_{ij} = \xi_{ij}^0 + \Delta\xi_{ij}$ where $\xi_{ij}^0 = 2r_{ij}/a$ while $\Delta\xi_{ij} = \epsilon_{ij}/kT$.

Let us split the percolation threshold into two parts,

$$\xi_c = \xi_c^0 + \Delta\xi_c \quad (6.10)$$

where ξ_c^0 is the known percolation threshold for the problem with bonding criterion

$$\xi_{ij}^0 \leq \xi \quad (6.11)$$

while $\Delta\xi_c$ is the correction which we are going to find.

In the case $K \ll 1$ let us first ignore the long-range potential. Then the most donor levels are close to the unperturbed value (which according to our notations corresponds to $\epsilon = 0$) and well separated from the value of chemical potential. Then for overwhelming majority of donor pairs we can neglect ϵ_i and ϵ_j in comparison with μ . As a result,

$$\Delta\xi_{ij} = \mu/kT, \quad (6.12)$$

and

$$\xi_c = \xi_c^0 + \mu/kT. \quad (6.13)$$

As a result,

$$\rho = \rho_{03} e^{\xi_c^0 + \mu/kT}, \quad \epsilon_3 = 0.99e^2 N^{1/3}/\kappa. \quad (6.14)$$

Since the long-range potential is small comparing to μ it produces only a small correction to the activation energy,

$$\epsilon_3 = \mu(1 - 0.29K^{1/4}). \quad (6.15)$$

6.3.2 High degree of compensation

Neglecting the long-range potential we can repeat all the estimates for the case $K \ll 1$ and set $\epsilon_{ij} = |\mu|$. In this way we arrive at the conclusion that

$$\epsilon_3 = |\mu| = \epsilon_D 2^{-1/3} (1 - K)^{-1/3}. \quad (6.16)$$

However, in most cases the long-range potential appears important. Since the long-range potential produces smooth distributions, in the vicinity of a point \mathbf{r} the donor energy is mainly determined by the energy

$$V(\mathbf{r}) = e\phi(\mathbf{r}). \quad (6.17)$$

The neighboring donors have almost the same energies (smooth potential!!). So, since $V(\mathbf{r}) \geq \mu$, we get

$$\rho(\mathbf{r}) = \rho_3 e^{[V(\mathbf{r}) - \mu]/kT}. \quad (6.18)$$

Now we arrive at the problem to calculate effective resistance of the system with an exponential spread of resistances, and we can apply the percolation theory for the second time.

Let us paint black the regions where $V(\mathbf{r})/kT \leq \xi$ leaving the remainder white. Suppose that only black regions are conducting while the white ones are insulating. Now begin increasing ξ , thus including the regions with higher resistivity. At some $\xi = \xi_c = V_c/kT$ the black regions merge into an infinite cluster, and further increase in ξ does not effectively change the resistivity. Therefore the resistivity is determined by the critical subnetwork and given by the expression

$$\rho = \rho_3 e^{(-\mu + V_c)/kT}, \quad \rightarrow \epsilon_3 = V_- \mu. \quad (6.19)$$

The quantity V_c can be found by numerical calculations. It is of the order of a typical fluctuation in the potential energy,

$$\epsilon_3 = C_1 \gamma(r_s) = C_1 \epsilon_D (1 - K)^{-1/3}, \quad C_1 \sim 1. \quad (6.20)$$

Form the above consideration we find the law

$$\epsilon_1 = \epsilon_3 + E_0 \quad (6.21)$$

which relates the activation energy to the conduction band to those in the impurity band. This relation has been confirmed by many experiments.

6.4 Variable range hopping

At low temperatures hopping occurs between the states whose energies are concentrated near the Fermi level. Because of the factor $\exp(\epsilon_{ij})/kT$ in the expression for the resistance R_{ij} , only the states within the narrow band near the Fermi level are important. The quantity ϵ_{ij} are given by Eq. (4.32). At sufficiently low temperature the density of states in this band can be regarded as constant, $g(\epsilon) = g(\mu)$, see Fig. 6.1. Similarly, we assume

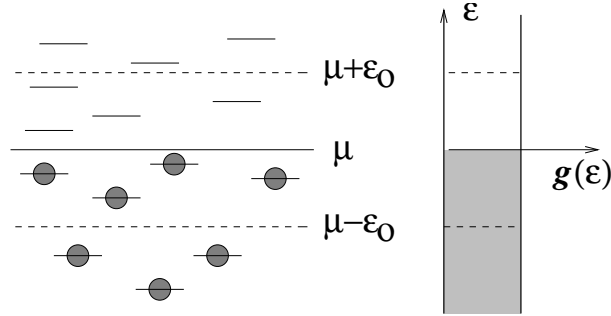


Figure 6.1: Construction of the band containing the states within the layer $\mu - \epsilon_0, \mu + \epsilon_0$.

that the localization radius is also energy-independent.

Let us consider the contribution of a small energy band

$$|\epsilon_i - \mu| \leq \epsilon_0, \quad (6.22)$$

which is symmetric about the Fermi level, see Fig. 6.1. The concentration of states in this band is

$$N(\epsilon_0) = 2g(\mu)\epsilon_0. \quad (6.23)$$

The typical resistance can be obtained by replacing the distance r_{ij} by the average one, $[N(\epsilon_0)]^{-1/3}$, and $\epsilon_{ij} \rightarrow \epsilon_0$. Then we have

$$\rho = \rho_0 \exp \left[\frac{1}{[N(\epsilon_0)]^{1/3} a} + \frac{\epsilon_0}{kT} \right] = \rho_0 \exp \left[\frac{1}{[g(\mu)\epsilon_0]^{1/3}} + \frac{\epsilon_0}{kT} \right]. \quad (6.24)$$

Here we ignore numerical factors. Let us analyze the dependence $\rho(\epsilon_0)$. It has a sharp minimum at

$$\epsilon_0 = \epsilon_0(T) = \frac{(kT)^{3/4}}{[g(\mu)a^3]^{1/4}}. \quad (6.25)$$

Substituting this value into Eq. (6.24) we get the famous Mott's law.

$$\rho(T) = \rho_0 e^{(T_0/T)^{1/4}}, \quad T_0 = \frac{\beta}{kg(\mu)a^3}, \quad \beta \sim 1. \quad (6.26)$$

The typical hopping length is then

$$\bar{r} = a(T_0/T)^{1/4} \quad (6.27)$$

which is temperature dependent. This is why the conductance of the type discussed above is called *the variable hopping conductance*. The coefficient β can be found with the help of percolation theory and numerical calculations. For 3D case $\beta \approx 21$.

For 2D case

$$\rho = \rho_0 e^{(\tilde{T}_0/T)^{1/2}}, \quad \tilde{T}_0 = \frac{\tilde{\beta}}{kg_2(\mu)a^2}, \quad \beta \approx 14. \quad (6.28)$$

Here g_2 is the two-dimensional density of states. Note that the derivation is based upon the assumption of a constant density of states at the Fermi level. We shall see that this assumption can be violated if electron-electron interaction is taken into account.

6.5 Coulomb gap in the density of states

At zero temperature, the distribution of electron is given by the minimum of electrostatic energy

$$\mathcal{H} = \frac{e^2}{\kappa} \left[\frac{1}{2} \sum_k^{\text{don}} \sum_{k' \neq k}^{\text{don}} \frac{(1-n_k)(1-n_{k'})}{r_{kk'}} - \sum_k^{\text{don}} \sum_i^{\text{acc}} \frac{(1-n_k)}{r_{ik}} + \frac{1}{2} \sum_i^{\text{acc}} \sum_{i' \neq i}^{\text{acc}} \frac{1}{r_{ii'}} \right], \quad (6.29)$$

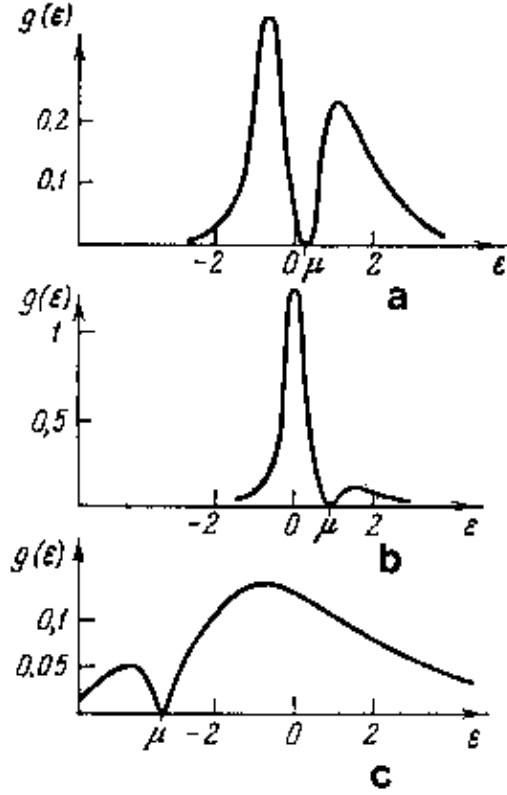
under condition that $\sum_k n_k = \text{const}$, or unconditional minimum of the functional

$$\tilde{\mathcal{H}} = \mathcal{H} - \mu \sum_k n_k, \quad (6.30)$$

where the Fermi energy μ should be then determined from the neutrality condition. We shall consider the energy counted from the chemical potential μ ,

$$\tilde{\epsilon}_i \equiv \epsilon_i - \mu = \frac{e^2}{\kappa} \left[- \sum_k^{\text{don}} \frac{(1-n_k)}{r_{ik}} + \sum_j^{\text{acc}} \frac{1}{r_{ij}} \right] - \mu. \quad (6.31)$$

The goal is to find the distributions $\{n_k\}$ and $\{\tilde{\epsilon}_k\}$ that is a complicated many-body problem. We have discussed is previously for the limiting cases $K \ll 1$ and $K \gg 1$. Now



Density of states in a lightly doped and compensated semiconductor in the presence of a Coulomb gap at different degrees of compensation. The energy ϵ is in units of $e^2 N_D^{1/3} \kappa^{-1}$, the density of states in $N_D (e^2 N_D^{1/3} \kappa^{-1})^{-1}$. (a) $K = 0.5$, (b) $K = 0.1$, (c) $K = 0.9$

Figure 6.2:

we turn to the case $K \sim 1$ and the only energy parameter is $\epsilon_D = e^2/\kappa r_D$, where r_D is the average distance between the donors. The results of numerical simulations of the density of states for different K is shown in Fig. 6.2. A gap at the Fermi level is clearly seen. The reason of this gap can be explained as follows.

Consider a band $(-\tilde{\epsilon}/2, +\tilde{\epsilon}/2)$ centered at the Fermi level. Let us estimate the energy cost of the electron transfer from the occupied state i to the empty one, j . Since the electron will be attracted to the donor, the energy cost is

$$\Delta_i^j = \tilde{\epsilon}_j \tilde{\epsilon}_i - e^2/\kappa r_{ij}. \quad (6.32)$$

It must be positive if the system was in the ground state. In other words, any two donors in this band having the energies in the opposite side of the Fermi level ($\tilde{\epsilon}_j \cdot \tilde{\epsilon}_i < 0$) must be separated in space by a distance

$$r_{ij} \geq e^2/\kappa \tilde{\epsilon}.$$

Consequently, the donor concentration, $n(\tilde{\epsilon})$ in a band of width $\tilde{\epsilon}$ cannot exceed $\tilde{\epsilon}^3 \kappa^3/e^6$. As a result, the density of states, $g(\tilde{\epsilon}) = dn(\tilde{\epsilon})/d\tilde{\epsilon}$ must vanish when $\tilde{\epsilon} \rightarrow 0$, at least as fast

as $\tilde{\epsilon}^2$. Because of that, the usually accepted formula for the density of states is

$$g(\tilde{\epsilon}) = \begin{cases} \alpha_3 \tilde{\epsilon}^2 \kappa^3 / e^6, & d = 3 \\ \alpha_2 |\tilde{\epsilon}| \kappa^2 / e^4, & d = 2 \end{cases} . \quad (6.33)$$

Here α_i are numerical coefficients, d is the dimensionality (in 2D case the density of states is calculated per unit area). It is important that the density of states has a *universal character*. A sort of self-consistent treatment provides $\alpha_2 = 2/\pi$, $\alpha_3 = 3/\pi$.

The Coulomb gap modifies the variable range hopping conductance. In 3D case the Mott's law is modified as

$$\sigma \propto \exp\left(-\sqrt{T_1/T}\right), \quad T_1 \approx 2.8e^2/\kappa a. \quad (6.34)$$

Chapter 7

AC conductance due to localized states

As we have seen, at low temperatures dc conductance vanishes in the insulating phase, the electronic states being localized. However, ac conductance remains finite. The reason is that an electron (or hole) can hop between the adjacent localized states giving rise to a finite energy dissipation.

Deeply in the insulating phase the most probable is to find only pairs of localized states which are separated by the distances much less than the average. To analyze the physics of dissipation, let us consider such a pair. We used such a concept to discuss Anderson localization.

Consider a close pair of localized states. As we have seen, the quantum states of an electron sharing the pair can be expressed through one-center wave functions as

$$\psi_{\pm} = c_1\phi_1 \pm c_2\phi_2 ,$$

which must satisfy matrix Schrödinger equation,

$$\begin{pmatrix} \Delta/2 - E & I \\ I^* & -\Delta/2 - E \end{pmatrix} \begin{pmatrix} \psi^+ \\ \psi^- \end{pmatrix} = 0 .$$

Here the origin for energy is taken at $(\epsilon_1 + \epsilon_2)/2$, while $\Delta \equiv \epsilon_1 - \epsilon_2$. The secular equation is thus

$$E^2 - (\Delta/2)^2 - |I|^2 = 0 \rightarrow E_{\pm} = \pm \sqrt{(\Delta/2)^2 + |I|^2} .$$

Consequently,

$$E_+ - E_- \equiv W = \sqrt{\Delta^2 + 4|I|^2} , \quad (7.1)$$

$$\frac{c_1}{c_2} = \frac{2I(\mathbf{r})}{\Delta \pm W} . \quad (7.2)$$

Thus at $\Delta \gg I$ either c_1 or c_2 is close to 1, and collectivization does not occur.

To find the dissipation one has to calculate the contribution of a single pair, and then sum the contributions of different pairs.

Dissipation by an isolated pair

One can discriminate between two absorption mechanisms. The first one is *resonant* absorption due to direct absorption of photons accompanied by transition between the states ψ^+ and ψ^- . The second mechanism is due to phonon-assistant transitions – an external field modulates the occupation numbers of close pairs. The modulation lags in phase comparing to the field due to relaxation. As a result, a finite dissipation appears. Below we consider both mechanisms.

Resonant contribution

The resonant absorption is due to transition of an electron from the state with the energy E^- to the state E^+ . The energy absorbed by a pair per unit time due to an electric field

$$\mathcal{E} = \mathcal{E}_0 \cos \omega t = \frac{1}{2} \mathcal{E}_0 [\exp(i\omega t) + \exp(-i\omega t)] \quad (7.3)$$

can be written as (Fermi golden rule!):

$$q = \frac{2\pi}{\hbar} \frac{\hbar\omega}{4} |e\mathcal{E}_0 \langle -|\mathbf{r}|+ \rangle|^2 \delta(\hbar\omega - W) (n_- - n_+) , \quad (7.4)$$

Occupation numbers

The occupation numbers n_{\pm} are determined from the following considerations. Let us write down a two-site Hamiltonian

$$H_{1,2} = \epsilon_1 n_1 + \epsilon_2 n_2 + \frac{e^2}{\kappa r} n_1 n_2 + I(r) (a_1^+ a_2 + a_2^+ a_1) . \quad (7.5)$$

Here n_i are the occupation numbers, while ϵ_i include Coulomb interaction with other neighbors.

The Hamiltonian (7.5) describes 4 states of the pair:

1. The pair has no electrons. The energy $E_0 = 0$.
2. The pair has 1 electron. There are two states with energies

$$E_1^{\pm} = \frac{\epsilon_1 + \epsilon_2}{2} \pm \frac{W}{2} .$$

3. The pair has 2 electrons. There is one state with the energy

$$E_2 = \epsilon_1 + \epsilon_2 + \frac{e^2}{\kappa r} .$$

Consequently, the probability to find a pair with 1 electron, the lower level to be occupied is

$$\begin{aligned} n_- &= \frac{1}{Z} \exp\left(-\frac{E_1^- - \mu}{kT}\right), \\ Z &= 1 + \exp\left(-\frac{E_1^- - \mu}{kT}\right) + \exp\left(-\frac{E_1^+ - \mu}{kT}\right) + \exp\left(-\frac{E_2 - 2\mu}{kT}\right). \end{aligned}$$

The occupation number $n_+ = n_- \exp(-W/kT)$, and we obtain finally,

$$q = \frac{2\pi \hbar \omega}{\hbar 4Z} |e\mathcal{E}_0 \langle -|\mathbf{r}|+ \rangle|^2 \delta(\hbar\omega - W) \exp\left(-\frac{E_1^- - \mu}{kT}\right) \left[1 - \exp\left(-\frac{\hbar\omega}{kT}\right)\right], \quad (7.6)$$

Relaxational contribution

To analyze this contribution one has to consider the balance equation, say for $n_+(t)$,

$$\frac{\partial n_+}{\partial t} = \frac{n_+ - n_0(t)}{\tau}. \quad (7.7)$$

Here

$$n_0(t) = \frac{1}{\exp[W(t)/kT] + 1}, \quad W(t) = \sqrt{[\Delta + e\mathcal{E}(t) \cdot \mathbf{r}]^2 + 4I(\mathbf{r})^2}, \quad (7.8)$$

while τ is the population relaxation time. Substituting

$$n_+(t) = n_0(t) + n_1(t), \quad n_1(t) \propto [e^{-i\omega t} + h.c.]$$

one easily obtains the relevant contribution to the absorbed energy as

$$q = \frac{\omega}{2\pi} \int_0^{2\pi/\omega} dt \dot{W}(t) n_1(t).$$

We get

$$q = \frac{|e\mathcal{E}_0 \cdot \mathbf{r}|^2}{2} \left(\frac{\Delta}{W}\right)^2 \frac{\omega^2 \tau(W, r)}{1 + [\omega \tau(W, r)]^2} \frac{1}{4kT \cosh(W/2kT)}. \quad (7.9)$$

The last factor is just $-(\partial n_0 / \partial W)$.

To calculate the dissipation one has to specify the relaxation time, which depends in general on W and \mathbf{r} . To do that let us specify the Hamiltonian to describe coupling between localized electrons and phonons. To construct the Hamiltonian, let us start with the unperturbed one,

$$\mathcal{H}_0 = \frac{1}{2} \begin{pmatrix} \Delta & 2I(\mathbf{r}) \\ 2I(\mathbf{r}) & -\Delta \end{pmatrix} = \frac{\Delta}{2} \sigma_3 + I(\mathbf{r}) \sigma_1. \quad (7.10)$$

Here we introduce Pauli matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Under influence of the phonon-induced strain the energy of each (j) component of the pair acquires the term proportional to the strain tensor,

$$u_{ik}(\mathbf{r}_j) = \frac{1}{2} \left(\frac{\partial u_i(\mathbf{r}_j)}{\partial x_k} + \frac{\partial u_k(\mathbf{r}_j)}{\partial x_i} \right).$$

Thus,

$$\tilde{\mathcal{H}}_{\text{int}} = \frac{1}{2} \sum_{ik} \left[\Lambda_{ik}^{(1)} u_{ik}(\mathbf{r}_1) - \Lambda_{ik}^{(2)} u_{ik}(\mathbf{r}_2) \right] \sigma_3, \quad (7.11)$$

where $\Lambda_{ik}^{(j)}$ are the component of the *deformational potential tensor* for each component.

Now we can make the transformation to a new basis which makes the Hamiltonian \mathcal{H}_0 diagonal. Technically, it is convenient to rewrite (7.11) in the form

$$\mathcal{H}_0 = \frac{W}{2} \begin{pmatrix} \cos \chi & \sin \chi \\ \sin \chi & -\cos \chi \end{pmatrix}, \quad \cos \chi \equiv \frac{\Delta}{W}.$$

Then one can represent the old basis through the new one as

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \hat{T} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}, \quad \hat{T} = \begin{pmatrix} \cos \chi & -\sin \chi \\ \sin \chi & \cos \chi \end{pmatrix} = \hat{1} \cos \chi - i \sigma_2 \sin \chi.$$

Having in mind the algebra for Pauli matrices,

$$\sigma_i^2 = 1, \quad \sigma_2 \sigma_3 = i \sigma_1, \quad \sigma_3 \sigma_1 = i \sigma_2, \quad \sigma_1 \sigma_2 = i \sigma_3$$

we obtain the interaction Hamiltonian in a new basis,

$$\mathcal{H}_{\text{int}} = \hat{T}^{-1} \tilde{\mathcal{H}}_{\text{int}} \hat{T} = \frac{1}{2} \sum_{ik} \left[\Lambda_{ik}^{(1)} u_{ik}(\mathbf{r}_1) - \Lambda_{ik}^{(2)} u_{ik}(\mathbf{r}_2) \right] \left(\frac{\Delta}{W} \sigma_3 - \frac{2I(\mathbf{r})}{W} \sigma_1 \right). \quad (7.12)$$

We are interested in the item proportional to σ_1 which is responsible for phonon-assisted transitions between the levels. Using Fermi golden rule to calculate the relaxation rate, we get

$$\frac{1}{\tau(W, r)} = \frac{1}{\tau_{\text{min}}(W)} \left(\frac{2I(r)}{W} \right)^2. \quad (7.13)$$

Here we have extracted the coordinate-dependent factor $(2I(r)/W)^2 = 2I_0 \exp(-2r/a)$. The quantity τ_{min} has a transparent physical meaning of the *minimal* relaxation time for a pair with given interlevel spacing W . It is dependent on several characteristics of the electron-phonon interaction which we do not discuss now.

Summation over the relevant pairs.

Now we can proceed adding the contributions of different pairs that appears rather tricky. What we need is to find the so-called *pair* distribution function which is the probability to find a pair with bare energy spacing W and spatial distance r .

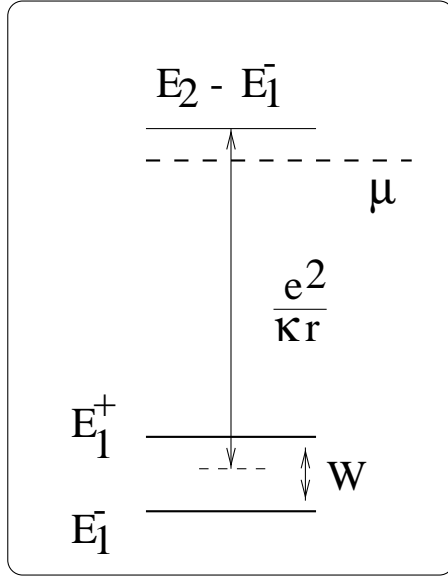


Figure 7.1: Energy scheme of absorption by a pair with the length r with allowance of Coulomb interaction.

Resonant contribution

For simplicity, let us assume low temperatures, $T \leq \hbar\omega/k$. Then $n_- - n^+ \approx 1$ at

$$E_1^- - \mu < 0, \quad E_1^- - \mu < E_2 - 2\mu.$$

It means that one can introduce the variables $\Delta = \epsilon_1 - \epsilon_2$ and $E_1^- = (\epsilon_1 + \epsilon_2 - W)/2$ instead of ϵ_1 and ϵ_2 . One can show that $\mathcal{P} \approx 1$ at

$$-\frac{e^2}{\kappa r} - W < E_1^- - \mu < 0,$$

and

$$\int dE_1^- (n_- = n_+) \dots \rightarrow \left(W + \frac{e^2}{\kappa r} \right) \dots$$

One can explain qualitatively this factor as follows from Fig. 7. When the center of gravity $(\epsilon_1 + \epsilon_2)/2$ falls into the region of the width $\sim e^2/\kappa r$ below the Fermi level the pair remains singly ionized and contains only 1 electron.

The full scheme of calculations can be described as follows. We define

$$\sigma(\omega) = \frac{2 \sum_i q_i}{\mathcal{E}_0^2} = \frac{1}{\mathcal{E}_0^2} \int d\mathbf{r} \int d\epsilon_2 g(\epsilon_2) \int_{\epsilon_2} d\epsilon_1 g(\epsilon_1) q(\epsilon_1, \epsilon_2, \mathbf{r}). \quad (7.14)$$

For the simplest (Anderson) model

$$g(\epsilon) = g_0 \Theta(A/2 - |\epsilon|).$$

Another important simplification arises from exponential decay of the overlap integral, $I(r) \propto \exp(-r/a)$. As a result, the coupling matrix element in the case of resonant contribution (as well as relaxation time in the case of the relaxational contribution) depend on r much stronger than other quantities. As a result, one can replace r apart from exponential functions by a proper characteristic length r_c (which is different for the resonant and the relaxational contributions).

Now we are ready to specify the results for low temperatures.. Substituting expression (7.15) for the contribution of a single pair we get (at $T = 0$)

$$\begin{aligned}\sigma(\omega) &= \frac{8\pi^2 e^2 g_0^2}{3\hbar} \int_{r_\omega}^{\infty} \left(\hbar\omega + \frac{e^2}{\kappa r} \right) \frac{r^4 I^2(r) dr}{[(\hbar\omega)^2 - 4I^2(r)]^{1/2}} \\ &= \frac{2\pi^2}{3} e^2 g_0^2 a \omega r_\omega^4 \left(\hbar\omega + \frac{e^2}{\kappa r_\omega} \right)\end{aligned}\quad (7.15)$$

where

$$r_\omega = a \ln(2I_0/\hbar\omega). \quad (7.16)$$

The case of finite temperatures is more complicated. At the region

$$\hbar\omega \ll kT \ll e^2/\kappa r_\omega$$

the width of integration over E_1^- remains essentially the same. However the occupation numbers of the states E_1^\pm become very close, and

$$(n_- - n^+) \approx \tanh(\hbar\omega/2kT).$$

This estimate is valid within all the region

$$\hbar\omega, kT \ll e^2/\kappa r_\omega.$$

The case of very large temperatures can be considered in a similar way. At

$$kT \gg e^2/\kappa r_\omega, \hbar\omega$$

both states E_1^\pm fall into the layer $\leq kT$ around the Fermi level, the difference $(n_- - n^+)$ being

$$(n_- - n^+) \approx \frac{1}{4 \cosh^2[(E_1^- - \mu)/2kT]}.$$

Thus.

$$\sigma(\omega) = \frac{2}{3} \pi^2 e^2 a g_0^2 \hbar \omega^2 r_\omega^4.$$

Relaxational contribution

Let us start with expression (7.9) which assumes that there are only the pairs with one electron. Introducing the distribution function $F(W, r)$ of the energy and spatial spacings (W and r) one obtains

$$\sigma(\omega) = \frac{e^2 \pi}{3kT} \omega^2 \int_0^\infty r^4 dr \int_0^\infty \frac{F(W, r) dW}{\cosh^2(W/2kT)} \frac{\tau}{1 + \omega^2 \tau^2}. \quad (7.17)$$

We will see later that $F(W, r)$ is a smooth function of r comparing to

$$\tau(W, r) \propto \exp(2r/a).$$

Thus let us define

$$\tau = \tau_0 \exp(2r/a)$$

where τ_0 is a smooth function of W and (in general) r . The properties of the ratio

$$G = \frac{\omega \tau}{1 + \omega^2 \tau^2} = \frac{\omega \tau_0 \exp(2r/a)}{1 + \omega^2 \tau_0^2 \exp(4r/a)}$$

depend strongly on the product $\omega \tau_0$. At

$$\omega \tau_0 \geq 1, \quad r \geq a$$

one has

$$G \approx \frac{1}{\omega \tau_0} \exp(-2r/a).$$

As a result,

$$\sigma(\omega) = \frac{e^2 \pi}{3kT} \int_0^\infty r^4 e^{-2r/a} dr \int_0^\infty \frac{F(W, r) dW}{\cosh^2(W/2kT)} \frac{1}{\tau_0(W, r)}. \quad (7.18)$$

Thus, σ appears ω -independent, the temperature dependence being determined by the properties of the functions $F(W, r)$ and $\tau_0(W, r)$. It is important that for the relevant pairs $W \sim kT$, $r = r_T \sim (a/2) \ln(2I_0/kT)$.

At

$$\omega \tau_0 \ll 1$$

the ratio G has a sharp maximum at

$$r = r_{c\omega} = (a/2) \ln(1/\omega \tau_0)$$

and we can express G as

$$G \approx \frac{4a}{\pi} \delta(r - r_{c\omega}).$$

In this way we obtain

$$\sigma(\omega) = \frac{4e^2}{3kT} \omega a r_{c\omega}^4 \int_0^\infty \frac{F(W, r_{c\omega}) dW}{\cosh^2(W/2kT)} \propto \omega. \quad (7.19)$$

Let us discuss a bit the distribution function F which is just the probability to find a *single-electron* pair with the energy distance W , spatial distance r . In the absence of Coulomb interaction,

$$F(W, r) = \int d\epsilon_1 d\epsilon_2 g(\epsilon_1)g(\epsilon_2) \delta\left(W - \sqrt{(\epsilon_1 - \epsilon_2)^2 + 4I^2(r)}\right) \\ \times \left\{ [f_0(E_1^-)[1 - f_0(E_1^+)] + f_0(E_1^+)(1 - f_0(E_1^-)) \right\}.$$

Here f_0 is the Fermi function. The integration is very simple for the Anderson model ($g = g_0 = \text{const}$). In that case

$$F(W, r) = g_0^2 \frac{[W^2 - 4I^2(r)]^{1/2}}{\tanh(W/2kT)} \sim g_0^2 kT.$$

Coulomb interaction changes the situation, and one has to calculate the number of pairs taking into account electron repulsion at the same center. At low enough $kT \ll e^2/\kappa r_{c\omega}$ that leads to the result similar to the one discussed above,

$$F(W, r) = g_0^2 \frac{[W^2 - 4I^2(r)]^{1/2}}{W} \left(W + \frac{e^2}{\kappa r} \right).$$

Thus at $kT \ll e^2/\kappa r_{c\omega}$ the absorption is essentially temperature-independent at $\omega\tau_0 \ll 1$.

Part II

Heavily Doped Semiconductors

Chapter 8

Interband light absorption

A typical experimental picture of frequency-dependent light absorption in heavily doped degenerate GaAs is shown in Fig. 8. The main features are:

- rapid decrease of absorption (4 orders of magnitude) inside the forbidden gap (1.51 eV at 77 K).
- Shift towards short waves with increasing concentration (Moss-Burstein shift).

The last feature is due to Pauli principle which requires that the transitions should take place only into empty states (above the Fermi level). For indirect transitions the shift is equal to μ while for direct ones it is $\mu(1 + m_e/m_h)$, see Fig. 8. This picture is relevant to very low temperatures. At finite temperatures there is an absorption below the threshold because one can find a hole in the conduction band with the energy $\mu - \epsilon$ with the probability $\exp(-\epsilon/kT)$. Thus,

$$\alpha \propto \exp[-(\mu + E_g \hbar\omega)/kT].$$

Another important source which we are going to discuss are transitions from *fluctuation levels* above the top of valence gap, Fig. 8. We see that the absorption is proportional to the probability to find a level with the energy $\epsilon_h = E_g + \mu - \hbar\omega$ which decays exponentially with ϵ_h . This probability must be multiplied by the transition probability which is not exponentially small in many important cases.

Thus, the frequency dependence of absorption in degenerate doped semiconductors reproduces the DOS profile for *minority carriers*.

The situation is somewhat different in the case of non-degenerate materials where the Fermi level is situated deep in the forbidden gap. The situation is realized at high temperatures or degrees of compensation, as well as in indirect materials.

Assume that there are impurities of both signs, $N_t = N_D + N_A$. The impurity potential is regarded as Coulombic at short distances $r \leq r_0$ and screened at large distances. Let r_0 be large enough to consider the potential as classical; we also neglect the correlation in defect positions. Mathematically, that can be expressed as

$$\frac{\hbar^2}{m_e r_0^2}, \frac{\hbar^2}{m_h r_0^2} \ll \gamma, \quad \gamma = \frac{e^2}{\kappa r_0} (N_t r_0^3)^{1/2}.$$

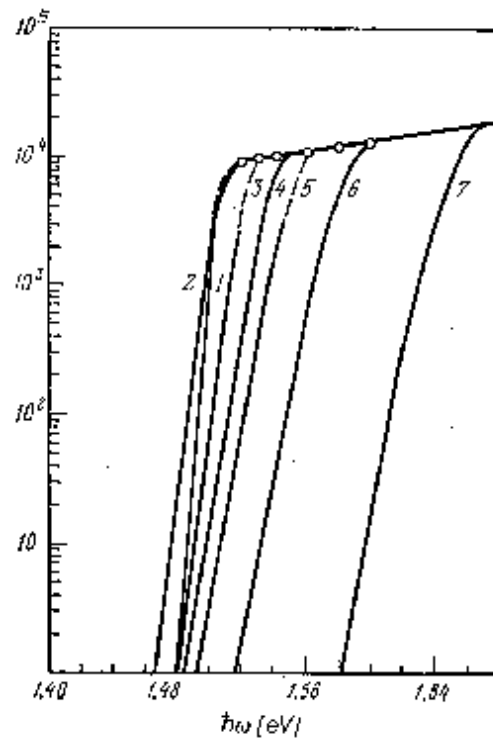


Figure 8.1: The absorption coefficient as a function of the energy of light quanta at 77 K in n -GaAs. $n \cdot 10^{-17}$, cm^{-3} : 1 – 0.02, 2 – 2.2, 3 – 5.3, 4 – 12, 5 – 16.2, 6 – 31.5, 6 – 65.

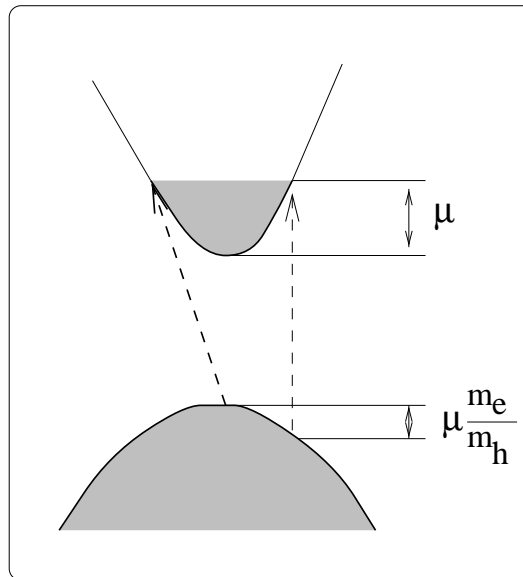


Figure 8.2: Scheme of typical interband absorption processes.

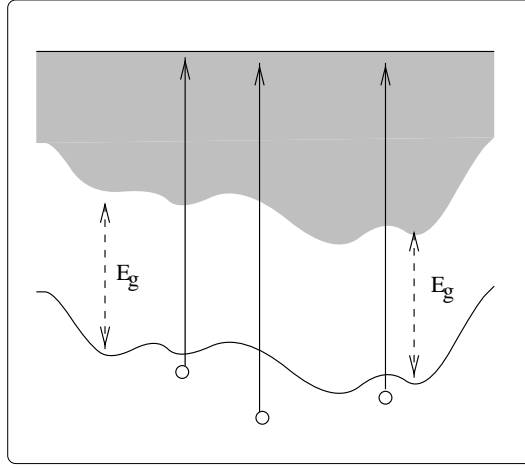


Figure 8.3: Scheme of interband transitions in a degenerate semiconductor at $T = 0$.

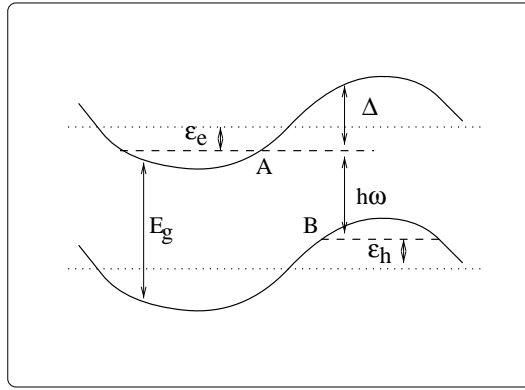


Figure 8.4: Absorption of a quantum of deficit Δ in a non-degenerate semiconductor.

One can show that in such a case the DOS tail can be represented as

$$g(\epsilon) = g(0) \exp(-\epsilon^2/\gamma^2).$$

In such an important case the absorption behavior *does not* represent the behavior of DOS, see Fig. 8. Suppose we are studying the transition which results in the formation of an electron of energy ϵ_e and a hole of energy ϵ_h ,

$$\Delta = \epsilon_h - \epsilon_e$$

(here the energies are calculated from the non-perturbed band edges). We observe that the electron and hole are *spatially separated*. Thus the matrix element involves *overlap integral* which has almost nothing to do with DOS. What we meet here, is the Franz-Keldysh effect in a random field.

To derive the result we use the conventional *optimum fluctuation method*. Here we demonstrate a simplified version of the method.

Consider a volume with linear dimension R . Charge fluctuations in the volume create a uniform electric field E , determined by the condition

$$eER = \Delta.$$

The excess number Z of charge defects to create the field $E = Ze^2/\kappa R^2$ must be

$$Z = \frac{E\kappa R^2}{e} = \frac{\kappa R\Delta}{e^2}. \quad (8.1)$$

The contribution of such a fluctuation to the absorption coefficient is proportional to

1. the probability to find Z excess charges in the given volume,

$$\exp(-Z^2/N_t R^2),$$

2. the probability for an electron to tunnel the distance R to meet the hole,

$$\exp\left(-R\sqrt{m_2\Delta}/\hbar\right)$$

(we assume $m_h \gg m_2$).

Substituting Z from Eq. (8.1) we obtain the probability to be proportional to

$$\exp\left(-\frac{\kappa^2\Delta^2}{e^2 N_t R} - \frac{R\sqrt{m_e\Delta}}{\hbar}\right).$$

Maximizing the exponent, we obtain

$$\tilde{R} = a \left(\frac{\Delta}{E_0}\right)^{1/2} \frac{1}{(N_t a^3)^{1/2}}.$$

Consequently,

$$\frac{\alpha(\Delta)}{\alpha(0)} = \exp\left[-\beta \left(\frac{\Delta}{E_0}\right)^{1/2} \frac{1}{(N_t a^3)^{1/2}}\right]$$

where β is the number of the order 1. The derivation is valid at

$$\tilde{R} \leq r_0.$$

In the opposite case, the probability to find the proper fluctuation decreases, and the optimal cluster has the size r_0 . In that situation one has to substitute r_0 instead of \tilde{R} to obtain

$$\frac{\alpha(\Delta)}{\alpha(0)} = \exp\left(-\frac{\Delta^2}{\gamma^2}\right).$$

Bibliography

- [1] B. I. Shklovskii and A. L. Efros, *Electron Properties of Doped Semiconductors*, Springer-Verlag, 1984.
- [2] C. W. J. Beenakker and H. van Houten, in “Solid State Physics” (H. Ehrenreich and D. Turnbull, eds.), v. 44, p. 1. Academic Press, Inc., 1991.
- [3] B. L. Altshuler, A. G. Aronov and B. Z. Spivak, JETP Lett. **33**, 94 (1981).
- [4] D. Yu. Sharvin and Yu. V. Sharvin, JETP Lett. **34**, 272 (1981)
- [5] D. E. Khmelnitskii, Physica **126b**, 235 (1984).